DOCUMENT RESUME

ED 436 963 FL 026 075

AUTHOR	Woytak, Lidia, Ed.
TITLE	Applied Language Learning, 1989-1998.
INSTITUTION	Defense Language Inst., Monterey, CA. Foreign Language Center.
ISSN	ISSN-1041-679X; ISSN-1041-6791
PUB DATE	1998-00-00
NOTE	1525p.; ISSN number varies.
AVAILABLE FROM	For full text: <http://lingnet.army.mil>.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Applied Language Learning; v1-9 1989-1998
EDRS PRICE	EDRS Price MF12 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Adults; Cognitive Measurement; Cognitive Style; Computer Assisted Instruction; Cooperative Learning; English (Second language); *Instructional Innovation; *Intensive Language Courses; Interlanguage; Language Laboratories; Language Research; Learning Strategies; Learning Theories; Second Language Instruction; Second Language Learning; Teaching Methods; Vocabulary Development

ABSTRACT

This journal is intended to be a vehicle for professional communication, discussion, and debate in the field of foreign language education, with a special emphasis on the instruction of adults in both commonly taught and less commonly taught languages, for functional purposes, in an intensive setting. Topics in this volume include the following: oral and written proficiency testing; second language acquisition research; teaching methods; the role of grammar in the development of communicative ability; training and use of new instructional technologies; the utility of objective measures of the fluency of speakers of English as a Second Language; instructional implications of gender differences in second language learning; sheltered initiation language learning; deviant English and the para-poetic; computer-assisted language learning in cooperative learning; reading versus grammar in second language learning; examining the value of language partner programs; pragmatics in interlanguage; sociolinguistic patterns of Arabic native speakers and the implications for teaching Arabic as a second language; validity issues in computer-assisted strategy assessment; the link between reading proficiency and native-like pauses in speaking; pedagogical intervention and pragmatic competence development; free voluntary reading as a predictor of TOEFL scores; an interactive model of vocabulary acquisition; pronunciation instruction for fossilized learners; and essay scores as instruments for placement and advancement in an intensive English program. Articles are extensively referenced. (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Applied Language Learning

Volumes 1-9

ED 436 963

FL026075

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL IN
MICROFICHE ONLY HAS BEEN GRANTED BY

Lydia Woytak

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

2B

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

VOLUME 1 · NUMBER 1

1989

DEFENSE LANGUAGE INSTITUTE FOREIGN LANGUAGE CENTER

APPLIED LANGUAGE LEARNING

Volume 1 • Number 1

1989

Defense Language Institute
Foreign Language Center

APPLIED LANGUAGE LEARNING

PB 65 - 89 - 1

Spring 1989 • Volume 1 • Number 1

Applied Language Learning (ISSN 1041-679X) is published semi-annually by the Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. *ALL* contains critical, integrative articles, and research reports on adult language learning for functional purposes.

COL Ronald I. Cowger, Acting Commandant
Defense Language Institute, Foreign Language Center
Department of the Army

This publication presents professional information, but the views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material to meet space constraints.

This publication approved for public release: distribution is unlimited.

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the Institute.

APPLIED LANGUAGE LEARNING

PB 65 - 89 - 1

Spring 1989 • Volume 1 • Number 1

EDITORIAL BOARD, DLIFLC

Neil Granoien, Editor-in-Chief	*Dean, Faculty and Staff Development*
Lidia Woytak, Executive Editor	*Faculty Development Specialist*
John L. D. Clark	*Dean of Evaluation and Standardization*
John A. Lett, Jr.	*Chief, Research and Evaluation*
Gary G. Walter	*Librarian*

EDITORIAL ADVISORY BOARD

David P. Benseler	*Editor, The Modern Language Journal*
James W. Dodge	*Middlebury College*
June K. Phillips	*Tennessee Foreign Language Institute*
A. Ronald Walton	*University of Maryland*

SPECIAL ACKNOWLEDGEMENTS

Annemarie Frazier
Word Processing Consultant

Virginia Wilson
Word Processing and Composition

APPLIED LANGUAGE LEARNING

PB 65 - 89 - 1

Spring 1989 • Volume 1 • Number 1

CONTENTS

vii From the Editor...

1 Features of Pragmatic and Sociolinguistic Competence in the Oral Proficiency Interview
Heidi Byrnes

15 The Role of Grammar in the Development of Communicative Ability
Nina Garrett

33 The Problem of the Target Model in Proficiency-Oriented Foreign Language Instruction
Albert Valdman

53 Testing Listening Comprehension
Dan Douglas

74 Information for Contributors

From the Editor . . .

In publishing this first issue of *Applied Language Learning*, one is tempted to make allusions to launchings, births, and the like. I shall resist seductive metaphors, however, and simply state that we at the Defense Language Institute Foreign Language Center are pleased to present what we anticipate will develop into a substantive and useful vehicle for professional communication, discussion, and debate in the field of foreign language education. The particular mission of our Institute is the instruction of adults in both commonly-taught and less commonly-taught languages, for functional purposes, and within an intensive setting. The Editorial Board will consider the scholarly articles and research reports of contributors that are relevant to those interests.

This issue contains four articles connected by a thread which weaves through topics of proficiency testing, second-language acquisition research, and pedagogy. Heidi Byrnes outlines possibilities for revising the oral interview proficiency guidelines to allow for a more detailed description of discourse competence that would focus on pragmatic and sociolinguistic indicators of fluency, a research area in need of further exploration. Nina Garrett addresses ambiguities in the description of the role of grammar in both language-specific proficiency guidelines and the teaching strategies they inform, and challenges foreign language educators to identify and develop the processing skills adult learners need to achieve grammatical competence. Working from data in sociolinguistic and psycholinguistic research, Albert Valdman posits a new native speaker construct for the target model of proficiency guidelines and shows its connection to pedagogical norms that can affect interlanguage development and fossilization. In his overview of recent directions in the testing of listening comprehension, Dan Douglas leaves us with important questions for validating a proficiency metric. How might we evaluate proficiency in terms of domains of behavior, rather than of discrete criteria? In what ways can we define the connection between context and testing? What effect might considerations of how language is learned have on the development of proficiency guidelines?

The Editorial Board would like to express its gratitude to those whose efforts have made *ALL* possible. Thanks go first to Joseph C. Hutchinson, who, before his retirement, laid the conceptual and procedural groundwork for the publication of the journal. Thanks are also in order to Lidia Woytak for the Herculean tasks she performed, although, with the fits and starts involved in any new enterprise, she might have felt the job more suitable for Sisyphus. We offer our thanks

as well as to Annemarie Frazier, whose technical expertise was of immense help, and to Virginia Wilson, who labored long, typing and editing multiple drafts. Finally, we want to extend our sincere appreciation to the members of the Editorial Advisory Board, David Benseler, James Dodge, June Phillips, and A. Ronald Walton, for generously sharing their experience and insights with us as well as for lending us their support and encouragement.

Applied Language Learning
Spring 1989, Vol.1, No. 1, pp. 1-13

Evidence for Discourse Competence In the Oral Proficiency Interview

Heidi Byrnes
Georgetown University

The paper examines the rating scale and the testing procedure of the oral interview for evidence of discourse competence. While the rating scale has generally been characterized in terms of function, content/context and accuracy, its underlying motivation derives from features of appropriateness and acceptability in message transmission. Given the limits of any testing procedure of oral communicative ability for capturing the multi-faceted pragmatics of language use, the paper suggests that a careful examination of the processing capabilities which underlie speakers' different ways of handling various types of meaning chunks or propositions might serve as a valid indicator of their discourse competence.

The evaluation of speaking proficiency according to the ACTFL rating scale, its testing protocol, the oral proficiency interview (OPI), and its procedure for assigning ratings to samples of speech have sparked a significant amount of controversy. One issue raised is that sociolinguistic and pragmatic parameters are inadequately represented while discrete-point factors receive a disproportionate weighting (Savignon, 1985).

The belief that ACTFL proficiency assessment places too much emphasis on the notion of accuracy, particularly grammatical accuracy, is understandable on two counts. First, it reflects a concern that prescriptive, static, and structuralist assumptions about L2 oral language use prevail in the OPI. Second, it challenges the assumption that oral language can be adequately assessed through a single testing procedure.

This paper is a revised version of a paper entitled *Features of Pragmatic and Sociolinguistic Competence in the Oral Proficiency Interview*, delivered at the Symposium on the Evaluation of Foreign Language Proficiency, held at Indiana University, March 4-6, 1987.

As for assumptions about L2 spoken language, it must be conceded that the pedagogical literature continues to presuppose an essentially static model of language, although an understanding of language as dynamic and interactive, and as embedded in a network of extralinguistic rules of language use is gradually being advocated. Even the research literature, though reflecting a more pragmatic bent than the pedagogical literature, often takes a restricted view. It tends to concentrate on the manifestation of specific L2 speech acts (e.g. Cohen & Olshtain, 1981) and rarely looks at global performance features that apply to discourse in general. Among such features, I count: rate of speech, silent and filled pauses, false starts, repetitions, corrections, stress and intonation features, all manifestations of what lay judgment has customarily referred to as fluency. In the context of assessing speaking ability the absence of research on fluency in its key manifestations of rate of speech and diversity of ways of constructing coherent discourse is particularly debilitating. Not only are important components of spoken language being ignored. More important for the present discussion, these components are precisely the ones that reflect speakers' discourse competence and thus, ultimately, their global language ability. Finally, and not surprisingly, almost no evidence exists for the evolving differences in discourse competence as learners acquire the L2 over an extended period of time and use it at different performance levels.

These circumstances in research make plausible the second concern about the OPI, namely, that raised from the testing perspective. Given our limited knowledge of the manifestations of pragmatic and sociolinguistic competence, except for an awareness of its dependency on a multitude of characteristics inherent in infinitely configured communicative situations, it seems inappropriate to assume that valid and reliable statements about L2 speakers' communicative ability can be generalized from a single testing procedure and a single rating hierarchy.

Cognizant of these concerns, I nevertheless suggest that we not dismiss the OPI prematurely but, as a first step, analyze the rating scale and the testing procedure from the perspective of interactive language use. Contrary to perception, this approach is clearly supported by the wording adopted in the revised *Guidelines* (Byrnes & Canale, 1986) and is overtly taken in the thoroughly updated tester manual currently being prepared under ACTFL sponsorship. I propose that the distinct gap in our awareness of L2 learner communicative ability, particularly discourse ability, can begin to be filled through the careful observation of performance as it is available through data from the OPI. We may well find that the bases for the rating scale are not unsound and that their articulation from the perspective of discourse

and pragmatics (Byrnes, 1987) is a valid approach toward better understanding of learner L2.

I will begin by addressing the extent and manner in which the current rating scale can and does incorporate discourse features. The result of this analysis will then be applied to the testing procedure. I will close with an overall assessment of that procedure and some suggestions for the direction future discussion might profitably take.

Of Hierarchies and Norms

The evaluation of oral proficiency under the ACTFL system proceeds from the stance that one can observe and differentiate language use along two axes: major types of speech events and different realizations of one and the same speech event. Such a position recognizes that any assessment of L2 language use must first refer to the message-conveyance tasks a speaker is able to perform and then to the degree of competence with which these tasks are being handled. Although it should be beyond debate, it is worth reiterating that this degree of competence is intimately tied to, though not exhaustively described by, a speaker's ability to access a diverse formal linguistic repertoire. Thus the real goal of the OPI is to determine a speaker's functional extra-linguistic competence through a well motivated extrapolation from the speakers' linguistic repertoire. There are noteworthy and well-known limitations to such a procedure. But, in the absence of other practicable and generalizable approaches, one should also acknowledge its significant potential for coming to grips, in a systematic way, with the pragmatic and social dimensions of language use.

Of the two axes mentioned above, the vertical one representing different global tasks reflects presumed native speaker judgments about the increasingly more elaborated ability, extra-linguistic as well as linguistic, required of speakers wishing to perform these tasks. In shorthand fashion, the progression of these tasks can be characterized as extending from naming and listing, to stating, narrating, describing, supporting an opinion, and negotiating or playing out a complex abstract issue or stance. The horizontal axis, which describes differences in the realization of one and the same speech event, does not primarily pertain to the absence or presence of errors in the traditional sense. Instead, it focuses on sociolinguistic appropriateness and the ways in which a communicative event is pragmatically enacted. It attempts to determine these dimensions on the basis of the perceived fit between what are taken to be the speaker's intentions and the overall form of the message.

Both axes derive their motivation from underlying concerns about the appropriateness and acceptability of a given speech sample *vis-à-vis* the norms of the speech community. It is imperative that norms be understood here not as distinct target behaviors, but as ranges of acceptable behaviors whose limits shift and adjust to numerous constituent factors of a speech event, among them setting, topic, intended message, and, of particular interest in our case, interlocutors. Native speakers, despite having essentially comparable command over language forms, possess all kinds of language performance profiles. In addition to such personality differences, L2 users show complex differences in the form inventories available to them. Although these diverse interlanguage profiles have tended to be interpreted primarily in terms of different degrees of formal, specifically grammatical, accuracy (see the debate about the article by Higgs & Clifford, 1982), the crux of the issue lies not in accuracy *per se* but in its constitutive role for meaning and the powerful metamessage it conveys about a speaker's discourse competence. It is this aspect and how it manifests itself in the OPI that I would like to explore in the following.

Overt and Covert Discourse Features

Let me begin by expanding on the task hierarchy from the viewpoint of discourse competence. This viewpoint essentially involves a shift from a focus on surface realizations–typically, the linguistic features which mark the word, sentence, paragraph, and extended text–to a focus on the underlying dynamics that lead to these surface realizations. These dynamics have at least two major inextricably interrelated components: the psycholinguistic processing characteristics of speakers and the stance speakers are able to take with respect to their message and their interlocutors. Both can be observed only indirectly, but powerfully nonetheless, under the rubric of what kinds of meaning chunks speakers can and do express in what kinds of language chunks. I take this relationship to be an important component of discourse competence and of pragmatic competence. Its particular value for articulating a hierarchy of L2 speaking competence into a viable frame of reference derives from its combining an internal developmental psycholinguistic progression with an external norm-oriented dimension.

As already emphasized, when we say that there exist different levels of speaking ability we are, among other things, saying that there exists a certain fit, based on a set of expectations, between communicative tasks and their linguistic realization. Instead of characterizing these realizations by their surface syntax, however, we might more profitably think of them in terms of ways a speaker has of

expressing global as well as specific meanings. In a conversational setting, these ways of expressing meanings are crucially related to speakers' ability to connect the individual propositions that manifest these meanings, both within turns as well as over extended instances of turn-taking. Such instances of cohesion and coherence in discourse reflect at once the speakers' own planning process as well as their concern with their listeners' ability to construct the intended meanings.

In examining how propositions are handled, we would find answers to questions such as these: Are propositions expressed primarily at the naming stage, which is closely tied to the lexical level? Can they be processed in simple relational terms, the essence of basic syntax? Is there evidence for the ability to hold a number of propositions simultaneously and to arrange these in a variety of increasingly complex ways, among them foregrounding, backgrounding, coordinating, subordinating, juxtaposing, attributing, assigning logical or chronological relationships, anticipating, referring back, etc.? Do the speakers exhibit the capability of selecting from among the major propositional relationships those that best serve their communicative intentions, or the demands of the communicative task at hand?

At first sight, one is tempted to reply that answers to these questions show little relationship to discourse as the term is commonly used, that they would primarily address linguistic phenomena, or that their real value would lie in pointing out retrieval difficulties of the learner. Or one might raise the concern that a distinction between speakers on the basis of how they realize propositions has more to do with their intellectual maturity than it does with their language competence. However, I would suggest that, while the interpretations mentioned are indeed appropriate, the progression implied by the previous questions is essentially one of discourse and pragmatic competence, in so far as the speakers are gradually enlarging their linguistic repertoire along a number of axes that address global language ability. Among these are axes that characterize language use as extending from situation-bound to context-independent, from reactive to initiating, from reactive.to overtly interactive to covertly interactive, from involved to detached, from speaker-bound to interlocutor-attentive, from uninvested to invested, from highly predictable task-specificity to individuated realizations of language-specific norms, among them norms of genre, style, register, or aesthetic valuation (see also Galloway, 1986).

In other words, I would see a relationship between the kinds of semantic processing speakers reveal through their language use and their discourse competence in that language (Garrett, 1986). When we say that somebody is a competent speaker of the language, we are not so much referring to the correctness or surface elaborateness that a speaker is able to mobilize for message transmission; we are really

referring to a host of discourse and pragmatic features. These features call for a speaker who:

(1) when necessary, is able to show, through language, independence of the immediacy of the situation; is able to follow the maxim of being communicatively cooperative and informative, as predicated by a particular setting;
(2) is able to take note of the wide variety of features inherent in a setting and seeks to make them responsive to her needs through language;
(3) has a level of awareness of the expectations of the interlocutors that allows him to frame his message to be recreatable and therefore meaningful for the hearers;
(4) formulates the message in such a fashion that various markers (hesitations, pauses, corrections, augmentations) and appropriate redundancies help the listener follow it easily;
(5) continuously monitors listener comprehension signals, actively seeking them should they be insufficient, and adjusting the message accordingly;
(6) uses discourse space in a way appropriate to the communicative task. Along with the ability to respond to the immediacy of communication, which usually results in lengthier and redundant speech, he can also plan discourse, a fact that reveals itself in reduced redundancies;
(7) makes the message compelling and dynamic, not merely because of overt give-and-take but because of a covertly interactive quality that arises out of anticipating the hearer's communicative expectations;
(8) is willing to take account of her audience, by delicately balancing positive and negative face (Goffman, 1967);
(9) is prepared to invest his ego into a message, rather than distancing himself from it (Galloway, 1986).

None of the above-mentioned ways of expressing propositions in itself carries a judgment of speaking proficiency; only in relation to global expectations for a wide range of communicative events, and with respect to refined expectations vis-à-vis a specific communicative event can such judgments be made. To give an example, it is perfectly appropriate to respond to a naming task by listing individual words. Consider, however, the following not at all infrequent occurrence with

second language speakers. A narrative task is handled in one or two initial minimal sentences. Interlocutor questions are required to elicit further minimal reactions, and each arises from a specific request by the listener that the narrator address a certain proposition which is assumed to have some import for the story. Depending on the patience, kindness, and level of interest of the listener/questioner the story actually might end up being told by the speaker. Even so, the result would be n instance of the non-native speaker not following the accepted rules of how a story is to be told, which could be due to any number of extra-linguistic factors, in themselves potentially weightier than the observed linguistic shortcomings. But in the absence of immediate analysis by the participants of their language use (a procedure that in no way guarantees the correct answer), we must make do with performance data. If such performance data analysis is backed by extensive correlational evidence, then assigning a lower rating is really an indication of the importance of pragmatic and sociolinguistic components in discourse. To rephrase, while the rating scale may, to a large extent, be expressed in terms of surface linguistic features, I believe, its real intent is to tap discourse competence through those surface language features.

We come here to two crucial issues in the assessment of speaking ability. First, in professing its basic feasibility, we must face up to the fact that while, for native speakers, linguistic options are the result of semantic options, the reverse may be true in the performance of language learners, particularly adult second language learners. Psycholinguistic and linguistic options may ultimately restrict the originally intended meanings and thus determine what meanings can be attempted in the L2, and the way in which they can be attempted (Canale & Swain, 1980; Bialystok, 1983).

Second, a rating scale will never be able to capture the reality of speech in all its facets (Bachman & Savignon, 1986). It will always, even if only tacitly, assume that evidence for performance feature *x* is grounds for extrapolating from it the speaker's ability to handle performance feature *y*. The key question is, what are the features that are considered baseline, and which features can, with some justification, be extrapolated? As it stands, the ACTFL rating scale represents the assumption that evidence for elaborated processing capabilities of the variety just described justifies extrapolation to the pragmatic and sociolinguistic features that contribute to discourse competence. This assumption finds initial support in interlanguage research (Bialystok, 1983). Since neither a rating scale nor a testing procedure can ever hope to capture, to an appreciable degree, the ability to create text, it is all the more important to explore and explicate this relationship.

Heidi Byrnes

Pragmatic and Sociolinguistic Features in the Oral Interview

When we speak of tapping discourse competence in the oral interview we are aware that, even with its focus on total language behavior, this procedure elicits no more than a minute sample of the virtually unlimited possibilities for speech events. The interview gives primary evidence for only one kind of oral language use, namely, the response a second-language speaker is able to provide to a single individual, the tester, in a setting that is marked by a power-relationship, whose intent is precisely that of testing oral language use. Little seems conducive in this setting for surfacing the crucial interactional, negotiative quality of speech; management features can only be tapped in a rudimentary fashion, and the function of language as a social bonder, or as the primary signaling system for the myriad of roles we play in our daily lives, to mention just a few aspects, simply cannot appreciably come into play. Even the restricted primary evidence of the interview, however, can yield a great deal of information if interpreted with regard to an aspect I alluded to previously, that is, the feature of fluency. Let us examine some properties as they are generally summed up by the term fluency.

The first set of properties broadly refers to characteristics of spoken discourse as contrasted with written discourse. Most important among these is the feature of spoken language being unplanned (Ochs, 1979) or, in view of the fact that all language involves planning, though at different levels, *impromptu* (Enkvist, 1982). Under the time pressures inherent in speech production, advanced editing is minimal, resulting in seemingly fragmented speech. The planning span is such that idea units are short and generally expressed in syntactically simple and lexically restricted fashion (Chafe, 1980). In addition, these idea units which are realized as intonation units involve numerous hesitations, false starts, repetitions, etc. and are themselves conjoined by simple chaining or coordinating. As a result, propositional density that requires planning is rather low.

In other words, if spoken language as a whole is compared with the literate norm, it is characterized more by disfluencies rather than by fluency, although we use that attribute in describing some speakers and withhold it from others. It seems, then, that fluency is, first of all, a relative concept and second, an interactive concept that depends just as much on the listeners' judgments as it does on the speaker's characteristics. Put another way, we can say that "fluent" speech is not so much flawless as it is disfluent in ways that are acceptable to the listener (Deese, 1980).

Given the crucial role of listener judgments in language assessment, information about such acceptable ways of being disfluent

takes on special significance. First studies exist to show how perceived fluency correlates with perceived higher levels of proficiency and which language features are considered crucial for both judgments (Schwartz, 1980; Gaskill,1980; Stewner-Manzanares, 1983; Kasper, 1985; Salo-Lee, 1989). It seems to be the case that the number of repairs does not necessarily decrease with increased proficiency. Instead, the range of types of repairs increases, and the more proficient speakers are those who are able to make the appropriate selections from this increased repertoire of repair strategies. Among these we count hesitations, restarts, clarifications, lexical searches, and correction invitations. Differentiated repair strategies serve both continuity, including the ability to fill the available discourse space, as well as discourse organization purposes.

While the summary descriptions of the major levels that make up the rating scale do not elaborate on such manifestations of fluency, the progression underlying the scale as well as actual rating practice clearly take them into account. For example, the ability to produce paragraph-length speech, a criterion for the advanced level, very often is manifested not in propositionally dense, lexically precise, and organizationally tight, polished paragraphs that we associate with literate language but rather in the extensive use of the kinds of repair strategies referred to above. It is, moreover, the speakers' ability to function as a cooperative partner with the interlocutor--for example, through propositional, even lexical repetitions in turn-taking instances --that creates the sense of a satisfying, harmonious exchange. More than specific instances of language use, then, it is the evidence of cooperative abilities that results in a higher rating.

Likewise, when an Intermediate level speaker is characterized as a separate speaker it is because hesitation and repair phenomena manifest themselves more on the local level and in looking back, as it were, over previous material, e.g., morphology, lexicon, and word order, rather than on the global level of cooperatively constructed discourse, which is often forward-looking. I would dispute the frequent interpretation that such speakers are characterized by a greater concern with accuracy or with form, than with meaning and successful message transmission. In reality, they, too, follow those presumably universal maxims. Due to processing constraints, however, their focus on message transmission manifests itself more in elaboration/correction on the local level rather than in larger meaning units. Furthermore, such speakers are judged as primarily reactive, since they tend to require a larger share of the work of cooperative meaning-construction from the listener, which results in their soliciting other-initiated and other-completed repairs, or other-initiated self-completed repairs. Thus, while fluency is in need of being spelled out in much

greater detail, the current rating scale would readily seem to allow such an emendation.

The Pragmatics of Testing

We have repeatedly referred to the interactive quality not only of the communicative event itself but also of its assessment. Of necessity, that realization puts a significant burden on the tester-rater, a burden that is made even more critical due to the highly restricted primary evidence obtained in the OPI. Nevertheless, I believe that, in the hands of a skillful tester, secondary evidence can become so powerful as to dramatically reduce the obvious pragmatic shortcomings of the setting and, in fact, permit some valid inferences.

Among many components that can be isolated, the key factor in skillful testing is the tester's desire to give and to receive real meaning rather than to contrive occasions that allow for mentally checking off whether the candidate can handle structure *x* or has available lexical item *y*. Neither of these can be a basis for assigning a rating. They are only a small part of the surface evidence we need for the underlying processing capabilities which, in turn, reveal pragmatic and sociolinguistic capabilities.

Time and again, I have observed that, when the tester is aware of conversational behavior–including listening behavior–and modifies it only so as to assure an optimal and efficient rating sample, the dynamics of message transmission become so powerful that the candidate can present an amazingly multi-faceted performance, even at the lower level of language use (Byrnes, 1986). And beyond the initial hurdles the imperatives of message transmission, coupled with the speaker's greater sociolinguistic competence, allow the restrictiveness of the setting to move further and further into the background and a true conversational exchange to take place. Conversely, when the tester does not bring to the task that level of interpersonal involvement and willingness to become a true conversational partner within the constraints of the event, little evidence for discourse competence can surface. An unratable sample is really the tester's inability to draw out of the candidate the best possible performance, an inability that is directly tied to a disregard of the pragmatic dynamics of conversational exchanges and of oral language use in general. One might interject that, at the lower levels, candidates are unlikely to be able to muster much pragmatic finesse anyway, and that, at the higher levels, they are quite capable of taking over. However, it is the in-between ground that is likely to suffer most under an inept tester, and, unfortunately, much of testing takes place here. To summarize, in order to permit the

clearest evidence for unperturbed language use, a genuine conversational exchange, deemed the archetypical occurrence of spoken language, is crucial for offsetting the substantial constraints of the testing procedure.

Outlook

Where are we to go from here? As indicated from my characterization of the rating scale, I would consider these points to be in need of further consideration:

> (1) The relationship between the primarily linguistic descriptors in the rating scale and the pragmatic and sociolinguistic parameters that underlie native-speaker judgments about non-native speakers' discourse competence, particularly at the more advanced performance levels (see also Albrechtsen et al., 1980).
> (2) The degree of acceptance, or tolerance, non-native speakers have for non-native speech. In contrast with discrete-point studies that have been undertaken, global assessments seem most instructive, particularly, if they are substantiated by an inventory and, if possible, a ranking of the criteria employed. On that basis, a greater awareness of the shifting norms for the appropriateness of non-native speech could be obtained.
> (3) Across the spectrum of second language research, the psycholinguistic and meaning-derived dynamics of language acquisition and language use must come even more into the foreground (for an example see Kramsch, 1983; Edmondson, 1985). For, in the end, teachers' perception of the second language acquisition process manifests itself in the language they present to learners and, consequently, the language learners learn.

While it may seem presumptuous to suggest that such a controversial topic as the proficiency rating scale and the evidence about speaking offered by its testing procedure, the oral interview, can have an important role in these inquiries, I would, in the absence of other viable contenders, nevertheless strongly urge its consideration.

References

ACTFL Proficiency Guidelines 1986 (1987). In H. Byrnes & M. Canale (Eds.), *Defining and developing proficiency: Guidelines, implementations and concepts* (pp. 15-24). Lincolnwood, IL: National Textbook.

Albrechtsen, D. Henriksen, B. & Faerch, C. (1980). Native speaker reactions to learners' spoken interlanguage. *Language Learning, 30*, 365-396.

Bachman, L. & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal, 70*, 380-390.

Bialystok, E. (1983). Some factors in the selection and implementation of communication strategies. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 100-118). New York: Longman.

Byrnes, H. (1987a). Speech as process. *Foreign Language Annals, 20*, 301-310.

Byrnes, H. (1987b). Second language acquisition: Insights from a proficiency orientation. In *Defining and developing proficiency* (pp. 107-131). See ACTFL Proficiency Guidelines (1987).

Canale, M.. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Chafe, W. (1980). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 169-180). The Hague: Mouton.

Cohen, A.D., & Olshtain, E. (1981). Developing a measure of sociocultural competence: The case of apology. *Language Learning, 31*, 113-134.

Deese, J. (1980). Pauses, prosody, and the demands of production in language. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 69-84). The Hague: Mouton.

Edmondson, W. J. (1985). Discourse worlds in the classroom and foreign language learning. *Studies in Second Language Acquisition, 7*, 159-68.

Enkvist, N. E. (Ed.). (1982). *Impromptu speech: A symposium*. Abo: The Research Institute of the Abo Akedemi Foundation.

Galloway, V. (1987). From defining to developing proficiency: A look at the decisions. In *Defining and developing proficiency* (pp. 25-73). See ACTFL Proficiency Guidelines (1987).

Garrett, N. (1986). The problem with grammar: What kind can the language learner use. *Modern Language Journal, 70*, 133-148.

Gaskill, W. H. (1980). Correction in native speaker-nonnative speaker conversation. In D. Larsen-Freeman (Ed.), *Discourse analysis in second language research* (pp. 125-137). Rowley, MA: Newbury House.

Goffman, E. (1967). *Interactional ritual*. Garden City, N. Y: Doubleday.

Higgs, T. V.. & Clifford, R. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57-79). Lincolnwood, IL: National Textbook.

Kasper, G. (1985). Repair in foreign language teaching. *Studies in Second Language Acquisition, 7*, 200-215.

Kramsch, C. J. (1983). Discourse function of grammar rules: Topic construction in German. *Modern Language Journal, 67*, 13-22.

Ochs, E. (1979). Planned and unplanned discourse. In T. Givón (Ed.), Discourse and syntax (pp. 51-80). New York: Academic Press.

Salo-Lee, L. (1989). [Investigating the repair strategies of learners of German at different levels of oral proficiency]. Ph.D. dissertation in preparation at Georgetown University, Washington, DC.

Savignon, S. J. (1985). Evaluation of communicative competence: The ACTFL Provisional Proficiency Guidelines. *Modern Language Journal, 69*, 129-134.

Schwartz, J. (1980). The negotiation of meaning: Repair in conversations between second language learners of English. In D. Larsen-Freeman (Ed.), *Discourse analysis in second language research* (pp. 138-153). Rowley, MA: Newbury House.

Stewner-Manzanares, G. (1983). Turn-taking structure of second language learners of Spanish at different oral proficiency levels. Unpublished Ph.D. dissertation. Georgetown University, Washington, DC.

Author

HEIDI BYRNES, Associate Professor, German, Georgetown University. Washington, DC 20057. Specializations: second language acquisition, testing.

Applied Language Learning
Spring 1989, Vol.1, No. 1, pp. 15-32

The Role of Grammar in the Development of Communicative Ability

Nina Garrett
University of Illinois, Urbana-Champaign

This paper suggests that to understand the relationship between grammatical competence and actual communicative language use we need a model of classroom language processing which considers learners' own connections between meaning and form as a major factor in language use and in language acquisition. This perspective suggests a different relationship between grammatical competence and communicative competence, presented here as an intuitive understanding of sociolinguistic, pragmatic and discourse appropriateness, not necessarily the ability to communicate appropriately. The paper considers the Proficiency Movement's (PM) representation of grammar in these terms, and argues that it is without psycholinguistic validity. It suggests that both L2 acquisition theory and FL pedagogy would benefit from development of the concept of grammar as processing rules in learning materials and in teacher training.

For the past fifteen years, the field of foreign language education has been increasingly concerned with developing a theory to include adequate representation of those other competencies beyond the grammatical, and with incorporating into practice techniques for inculcating them in language learners. A heavy emphasis on communicative competence has been, and continues to be, crucial in correcting the traditional view that language consists only of a lexicon and a set of

---

An earlier version of this paper was given in March, 1987, at Indiana University at the Symposium on the Evaluation of Foreign Language Proficiency. I am indebted to all those who spoke at the Symposium for the clarification of issues, sharpening of focus, questions and criticisms which contributed to the final version of this paper. I am most particularly grateful to Albert Valdman for invaluable advice and encouragement. In addition I am indebted to Gary Cziko, John Lett, Alice Omaggio, and Bill VanPatten for their comments on earlier versions. Any inadequacies or heresies remaining are my responsibility.

rules for combining words into correct sentences. That the focus of this paper is on grammar does not imply any reservation about the importance of the other competencies, nor any intent to regress to a methodological position asserting the centrality of traditional grammar. On the contrary, my focus is motivated by the concern that in our zeal to realize the implications of the construct of communicative competence we have allowed the notions of grammar and grammatical competence to slip into an uneasy limbo. Although no one denies that communicative ability must somehow include grammaticality, there is no general agreement in the field as to whether or how the *facts* of grammar as they have always been conceived can be taught so as to transmute into the ability to speak grammatically. In other words, no principled relationship between grammatical competence and communicative ability has yet been successfully put forth. The postulation of such a relationship requires a complete revision of current notions of what grammar is and how it operates in language use. This paper, therefore, addresses the theoretical problems underlying the controversy over the role of grammar not only in the PM but in the field of foreign language education generally.

Grammar in Testing and Teaching for Proficiency

The debates over the adequacy of the PM's representation of communicative competencies (pragmatic, sociolinguistic, and discourse) have concentrated on three questions: (a) whether the *Guidelines* are based on a sufficiently sophisticated concept of the particular competence; (b) whether the OPI can elicit a language sample adequate to give ratable evidence of that competence; and (c) whether a syllabus based on the *Guidelines* can inculcate it. The argument, in short, has been whether the PM represents an adequate understanding of these concepts and steers closely enough toward them.

In contrast, the debate over the PM's representation of grammatical competence has focused on the questions: (a) how much it should be weighted in the testing of global proficiency, and (b) whether there are either theoretical or empirical grounds for allotting much (if any) classroom time to teaching grammar. It is taken for granted that the concept of grammatical competence is understood; the argument has been whether the PM steers far enough away from it.

The Issue of Grammar in the Field at Large

In fact, this debate is not new to the field. The question of whether or not grammar should be explicitly presented has plagued the field for

centuries (cf. Kelly, 1969) and has been a central issue in debates about communicative competence since that concept became an organizing principle for foreign language education some 15 years ago. Almost everyone pays lip service to the claim that grammatical competence is an essential part of communicative ability, but there is no generally accepted theoretical or methodological stance on how it is to be achieved.[1] The arguments over grammar are usually either that there is little reason to believe that (a) teaching grammar does any good and much reason to believe that it may be harmful, or (b) not teaching grammar is better and some reason to believe that not teaching it may be harmful. And each side attacks the research cited by the other. The PM includes "accuracy" as one of the three major components of communicative ability and is generally seen as promoting attention to grammar, but there seems to be a wide range of opinion as to the desirability of teaching grammar explicitly.[2]

Underlying the debate, however, both camps–those who believe that teaching grammar is necessary for classroom second language acquisition and those who believe that it is not only not necessary but generally not even advisable–hold to strikingly similar fundamental assumptions that:

> (1) the nature and function of grammatical competence is established and understood;
> (2) to communicate grammatically one must somehow acquire the rules of competence, but that one can and should do so without too much overt attention;
> (3) the nature of communicative competence is generally understood, though the particulars are still being developed;
> (4) communicative competence is more important, more interesting, and requires far more attention than grammatical competence.

None of these assumptions should be taken for granted; together they form an inappropriate set of preconceptions about what grammar is and about the role it plays in classroom second language acquisition.

The Basis for the Role of Grammar in the Guidelines

The *Guidelines* as rating level descriptions were originally developed on the basis of experiential data, reflecting the job-related language requirements of the population being tested, rather than on the basis of adult classroom second language acquisition theory or theoretically motivated empirical research. The specification of the

sequence in which certain structures are expected to be produced is dependent on a more or less arbitrary or *intuitive* ordering of the functions, and expectations for the development of accuracy are based on the production of large numbers of learners. This needs no defense; experiential data constitute an altogether reasonable starting point for developing such a test. As long as the OPI is being used simply as a test of certain aspects of what we teach, it need only have reliability and face validity; as long as those who use it agree that the test results correlate to their satisfaction with learners' ability to function in the FL under specified conditions and for specified purposes, it might be argued that no independent theoretical basis is required for the specifications of accuracy.

Again, if the *Guidelines* are to be used to guide syllabus design, the crucial question is the justification claimed for doing so. We are certainly at liberty to decide that we want our students to be able to carry out certain functions in certain contexts, and to specify the sequence in which we want to present these functions and contexts explicitly in class. If we want students to be able to translate, we design a syllabus to inculcate that ability; if we want them to carry out certain communicative functions, we design a functional syllabus. If we agree on the specifications for a syllabus, it needs no theoretical basis. Deciding on a syllabus is a matter of pedagogical policy.

Current developments in the PM, however, seem to be based on a serious confusion of pedagogical policy with theory. If the *Guidelines* are used to motivate syllabus design and teaching practice not only because they furnish a functional description of a specified pedagogical goal but also because they are believed to describe how adults in the classroom learn a foreign language, then a theoretical claim is implied which requires independent support. Without a coherent theoretical basis, recommendations for syllabus design and teaching practice based on experiential data are inevitably circular. They imply that because most learners in the past seem to have learned this way, this must be the natural way for them to learn and therefore this is the way they should be taught. This circularity becomes genuinely dangerous when it motivates suggestions to the effect that since the *Guidelines* indicate that most learners do not seem to be able to use grammar very well until later levels, maybe there is no point teaching grammar until they get to those levels.

It might be argued that most proponents of the PM do not in fact make any theoretical claims, and it is true that the field of foreign language education as a whole tends to pride itself on being atheoretical, on its eclecticism, on "caring about what works." But the apparent unresolvability of the arguments about the value of explicit grammar indicates that we often do not know what works, and we cannot undertake research on what works regarding grammar without

developing a coherent and testable theory of its role in the development of communicative ability. Partly because grammar has become a fighting word, and partly because we have always assumed that we knew what it was, no one has wanted to undertake to develop such a theory, but the momentum gathered by the PM makes it absolutely necessary. Without a theory of how grammar works, the PM has no substantive basis for claiming that the OPI assesses grammatical competence in any useful sense, still less a basis for claiming that the *Guidelines* provide any guide for whether, or how, to teach it.

The problems connected with assessing the role of grammar in the PM are not so different from the problems attending the other components of communicative ability. The questions about the role of grammar should not focus on how we can downplay it but should be of the same form as the other debates, i.e., whether the concept of grammar which informs the *Guidelines* is adequate to serve as one of the structural members of a framework for syllabus design or to assure that the OPI can elicit data which will allow us to assess, in any interesting way, the grammatical competence of our learners. For the moment the answer must be negative.

Theoretical Assumptions Underlying General Research on the Role of Grammar

Research on the role of grammar in foreign language education and natural second language acquisition has been for the most part limited to studies of how learners *use* those rules of morphology and syntax which constitute the traditional linguistic description of a language. Addressed as well are the ways learners acquire those rules (whether they need to be taught explicitly, or whether explicit learning gets in the way of automatic use), what makes the rules accessible for use, and under what circumstances learners can access them.

Those morphological and syntactic rules are generally said to constitute the *competence* of Chomskyan theory, but strictly speaking competence does not comprise the rules but only the intuitions of native speakers about how the grammatical structures of their language are related to each other in a system. It is from these intuitions that the abstract rules or generalizations are then derived by linguists. The idea that competence, the underlying knowledge of rules, is the basis for communicative language performance is widely accepted. The concept of communicative competence is the essential further bit of theory representing the notion that there is more to speaking a language than implementing grammatical knowledge. The rules of communicative competence are commonly characterized as *telling us how to use* the rules of grammatical competence, as mediating between it and

performance.[3] Thus the rough general model of language acquisition has competence as the central construct in second language acquisition theory; competence is thought of as *that-which-is-acquired* because it is thought of as that-which-is-used as the basis for performance, and performance is the goal we have explicitly insisted on ever since the demise of the grammar-translation method. We want our students not merely to know about the language but to be able to communicate with it.

The basic theoretical argument of this paper is that allowing competence to occupy a central position has confused our attempts to develop a coherent theory of the appropriate role for grammar in classroom second language acquisition. In any psychologically real sense linguistic rules are not the basis on which anyone, native speaker or learner, actually produces or comprehends an utterance. Still, the unarguable goal of language learning is that the learner should come to be able to produce and comprehend utterances which fit the linguistic description of the language being learned; it cannot be suggested that the linguistic notion of a grammar has no role to play in accounting for language use or language acquisition. How then do grammatical intuitions or rules really operate in the mind of a speaker?

The Relationship Between Meaning and Grammatical Rules

In producing or comprehending an utterance a speaker-hearer does not start with competence. The basis for language behavior is not a set of intuitions about how a language system as a whole works; the basis for performance is the desire to express some meaning. Meaning is encoded into utterances through a complex set of processing rules, psycholinguistic rules which operate in real time in speakers' minds to combine lexical items and a variety of formal features (word order, inflections, etc.).[4] The linguistic facts are psycholinguistically relevant only as they are present in the mind of speakers, that is, as they are used in processing. If, however, intuitions of grammaticality are not the basis or starting-point for processing, how are they used? Psycholinguistic plausibility suggests that they constitute instead the end-point, that competence can best be understood as a template, a sort of Platonic ideal of the particular language, against which any utterance has to be mentally matched.

A speaker begins to construct an utterance on the basis of meaning. The various minute and subtle bits of the intended meaning take form in lexical items and formal features or grammar. (Every language has a set of form types: word order, word changes or morphology, stress, tone, etc., all of which are included in this view of grammar). Words and forms are combined in an utterance whose configuration must fit the

speaker's intuitions about the linguistic system, the template of grammatical competence.

Meaning, however, is not limited to semantic content; pragmatic, sociolinguistic, and discourse meaning are integral to every adult utterance. Communicative competence is then the set of our intuitions about how these kinds of meaning are also represented in words and formal features. An utterance can be grammatical without being communicatively appropriate, but it cannot be appropriate unless it is grammatical. Communicative competence, therefore, includes grammatical competence but goes beyond it. Communicative competence in this sense is thus exactly analogous to linguistic competence: it is not an ability, but rather refers to another kind of knowledge, another set of intuitions, above those which confirm grammaticality, and thus it functions as another template, a set of supra-grammatical constraints on language behavior.

The linguistic rules of a language are derived from native speakers' intuitions about what is grammatical; the discourse, sociolinguistic, and pragmatic rules of a language are derived from native speakers' intuitions about what is appropriate. All these rules have psycholinguistic validity, in that they function as a body of knowledge against which the result of production or comprehension can be checked; but they should not, however, be thought of as a starting point for the actual production or comprehension of utterances. Nor–and this is the crucial point–can they be thought of as the rules according to which production and comprehension take place.

The *rules* which are used in any psycholinguistically real sense in actual speaking and understanding are the patterns of association by which words and formal features are attached to bits of meaning (all kinds of meaning) to be conveyed or comprehended. These are often called processing rules, since they process meaning into language form (in production) and process form into meaning (in comprehension).

The Teaching of Grammar Re-Examined from the Processing Perspective

Once we see processing as the central construct in actual communicative language use, with both kinds of competence relegated to the role of final template, it is easy to understand why the explicit teaching of linguistic rules does not contribute much to learners' ability to communicate grammatically. Linguistic rules–the abstract formulations describing how the language system works–are not per se processing rules. Explaining the linguistic template (telling learners what rules their finished utterance must conform to), therefore, will not of itself tell them how to construct for themselves the associations, the

processing rules, which connect meaning to form in any particular language.

However, because the two sets of rules are related, language learners can often work from one set to the other. Even in classes which emphasize old-fashioned grammar instruction, some learners can develop communicative ability, which means that they can extrapolate processing rules from linguistic rules. (That is, they can work out meaning-form connections from discussions about the abstract relationships between formal features.) In contrast, in classes where the presentation of linguistic rules is sharply reduced or avoided, the *good* learners are those who can extrapolate those meaning-form connections from input instead. Since input is not always directly or unambiguously meaningful, the learner is, in effect, being asked to hypothesize a connection between form and meaning without either being specified. Learners differ in their preference for inductive versus deductive learning, and it may be that those who prefer the former will acquire communicative ability better without explicit grammar than those who prefer deductive strategies. The point is that the processing perspective allows us to ask an entirely different question: would it help students if the processing rules were explicitly explained, so that they would not be forced either to induce or to deduce them?

In the short run, what seems to help the students may be the simple surface structure collocations and heuristics which are virtually empty of meaning, such as the all-too-familiar textbook grammar pointers that tell them *If there's an X in the sentence, use a Y here.* But, in the long run, any syllabus or method that aims to enable learners to do something with the language, rather than only to know something about it, has to make some claim (even if only implicitly) about all the components of language performance: the meaning, the formal features, and the processing which connects them. Different methods or approaches will, of course, place their primary focus on different components. In terms of the processing model proposed here, the grammar-translation method focused on the rules in the linguistic template (although it did not pretend to teach communicative ability). Communicative approaches focus on sociolinguistic, pragmatic, and discourse meaning, avoiding excessive attention to the linguistic template (although different versions of the communicative approach will differ as to what is considered excessive). Some variants of a communicative approach regard sociolinguistic meaning as primary, while a functional approach often places special emphasis on the pragmatic. A notional approach in principle stresses semantics, but the notional functional syllabus tends in practice to be mostly functional, if not outright situational. The crucial point is that every authentic adult utterance represents all kinds of meaning, and so only an understanding of the

processing that goes into the representation of all kinds can form an appropriate theoretical basis for our teaching.

The actual classroom practice of most teachers is usually based less on any theoretical position than on a set of hunches (or preconceptions based on their own language learning experience) about what seems to work, even if they are not trained or inclined to think about the form-meaning connections involved in basic grammatical notions. The capsule descriptions of syllabus philosophies given above thus leave out of account the practice of teachers who reject labels and party lines and work out their own compromise position on grammar. Some, for example, feel the need for a grammatically structured syllabus but seldom if ever explain a grammatical structure explicitly for its own sake, preferring to focus students' attention on the communicative import of a structure by highlighting its use in communicative contexts. Others reject a grammatically structured syllabus in favor of a functional one, but may provide explicit explanation and drill of whatever grammatical structure arises in connection with a given function, perhaps providing it entirely outside of classroom time via workbook or computer. Many other compromise positions can be imagined and are undoubtedly practiced by experienced and successful teachers, and their intuitions about what works cannot be chopped or stretched to fit any Procrustean bed of pedagogical theory about explicit grammar.

There are obviously teachers who have developed a variety of ways to present linguistic rules, so that learners can readily translate them into processing rules; and obviously there are learners who have little difficulty working from linguistic rules to develop communication strategies and appropriate early-interlanguage rules. We have always known that good language learners can learn well, whatever the approach; and we have always known that good language teachers can teach well, regardless of their theoretical or methodological training. There is no guarantee, however, that any teacher will with experience develop the requisite insights and techniques. The concept of processing provides us with a principled basis for understanding what we already know intuitively about good learning and good teaching, from which we can generalize to help less good teachers and learners.

The concept of processing stresses the close relationship which exists between the linguistic and the processing rules, and makes clear how it is possible for some learners to work from one to the other. Our textbooks do explain some linguistic rules in terms of the form-meaning relationships they represent. Most conventional textbook explanations of linguistic rules, however, are couched either in grammatical terminology (under the assumption that terms like *direct object* or *tense* are meaningful to the learner), or in terms of correspondences to the meaning allegedly conveyed by a given structure in English, or in terms

of correspondences with another structure (for example *explaining* the passive in terms of its relationship with the active). For a great many of our students these explanations cannot, in fact, be assumed to convey any concept whatever. Furthermore, even when a form is explained immediately in terms of some bit of meaning, the value of the explanation is frequently nullified by explicit or implicit indications of surface structure clues by which the learner can guess what form is called for without having to analyze meaning. As a result, learners develop, as interlanguage rules, heuristics for using TL forms to convey an element of meaning quite different from that with which the TL actually marks them.[5]

What would be the implication for our teaching practice of a more fully developed theory of grammar-as-processing-rule? The most immediate is the importance of recognizing the integrity of processing, the integrity of communicative language ability. The separation of form and meaning is not only impossible in terms of all that we know about the nature of language knowledge and language use but also counter-productive in terms of our philosophical goal of helping students express their own meaning in the appropriate forms of the target language. The separation of grammatical and communicative competence is equally impossible: meaning of any kind--semantic, relational, propositional, sociolinguistic, pragmatic, discourse--can only be conveyed through appropriate use of grammatical form, at least beyond the single-word level. Nor can the processing of the lexicon be separated, except in superficial terms, from the processing of grammatical form; there is no fundamental difference between learning how lexical items convey meaning and learning how bits of form do so.

But this discussion of the importance of conceiving of grammar in terms of the processing model should not be taken to imply that processing rules can be handed to students on a silver platter, that explaining grammar in terms of the form-meaning connections will instantaneously enable learners to speak grammatically. Processing cannot be directly taught to guarantee production that is either grammatically accurate or communicatively appropriate, any more than grammaticality can be taught just by explaining linguistic rules. Nonetheless, the concept of processing itself is surely crucial as a framework within which learners can organize the language data they get from communicative input or any other source. Beyond presenting that framework, three steps in "teaching" processing rules which are realized in grammatical form could be suggested: a) making students aware of a particular bit of meaning (of whatever kind) which must be conveyed grammatically in the TL; b) setting up the form-meaning connection at whatever level of refinement is appropriate for the language-learning level; c) providing ample opportunity to practice the association, including communicative exercises, when the accurate

representation of meaning requires syntagmatic as well as paradigmatic choices.

Notice that *making aware* and *setting up* and *providing opportunity* are all deliberately vague. In each language, for each form-meaning connection, decisions about what degree of explicitness is needed should be motivated by research on the adequacy of the actual processing rules which students are likely to make up for themselves in the course of acquisition; the more idiosyncratic and inadequate these are, the more explicit the explanation should be. The traditional assumptions held by teachers and textbook authors about the transparency of form-meaning connections and students' ability to induce or create them are often wrong. Because we have barely begun to study learners' interlanguages from this point of view we do not yet know which early crude interlanguage rules are likely to be useful and inherently refinable as opposed to those which are likely to be dead-end and a permanent hindrance to the construction of a TL-like processing rule.

The global implication for foreign language education of understanding grammar in terms of processing is that we must change our efforts to focus on learning rather than on teaching, and we must address the individual learner's needs. There is no best method or best approach for all classroom foreign language learners. Differences in individual learners' affective characteristics have been recognized for years now, and differences both in cognitive style and learning style are receiving increased attention. The realization that individual learners' creation of interlanguage processing rules is also highly idiosyncratic makes inescapable the conclusion that the only viable approach to grammar is (1) to help the students themselves to think about language acquisition and language use in processing terms so that they can understand their own learning needs and predilections, and (2) to maximize the individualization of learning options for them.[6]

The Problem of Sequencing

Even with a theoretical framework motivating the explanation of grammatical structures as coding elements for the processing of meaning, we still need principles for deciding on a sequence of presentation of those structures. Neither child acquisition orders or those of natural learners can be taken a priori to be appropriate for the classroom. Is there any theoretical criterion for structural sequencing?

One approach is suggested by the work of Talmy Givón (1979), who insists that language can only be understood from a functional perspective, a claim which should recommend his work to us. He posits two extreme poles on the continuum of native speakers' communicative

language use, the pragmatic mode and the syntactic mode, and explicates a number of different relationships between them. The pragmatic mode is characterized as having topic-comment structure, loose conjunction, slow rate of delivery, pragmatically governed word order, and little or no use of grammatical morphology, in contrast to the syntactic mode's subject-predicate structure, tight subordination, fast rate of delivery, word order governed by semantic case function, and elaborate grammatical morphology. Givón argues that the contrast between the pragmatic and the syntactic obtains not only diachronically in the development of a given language (and to a certain extent his arguments on this point fit in with Slobin's explication of "Language change in childhood and in history," 1977) but also between pidgins and creoles, between adult and child language, and between informal/unplanned speech and formal/planned language. He explicitly mentions the relevance of this distinction to the adult second language learner: "Under a pressing need to communicate in a foreign language, the adult does, in fact, revert almost all the way to the extreme pragmatic register on the scale."

To a remarkable extent Givón's arguments overlap with those made by Corder (1977). The simple code which Corder postulates is clearly much the same as Givón's *pragmatic mode*, and indeed Givón's terminology solves the problem, which Corder addresses at some length, of the undesirable connotations of *simple* and *simplified*. Many of Valdman's (1980) recommendations for a modified structural syllabus are based on similar analyses of the relationship between pragmatically motivated paratactic constructions and the syntacticized ones mandated by traditional textbook presentations. The similarity of these claims and of their attempts to account for quite different problems and data (although they use different terminologies in their discussion) suggests that this view of discourse could form a significant connection between linguistics, second language acquisition, and foreign language education, and could provide us with testable hypotheses about the order in which syntactic and morphological features appear in learner language.

Givón's theoretical framework is particularly suitable for our current priorities in that he insists on the primacy of discourse and the need to consider grammatical rules as an automatic processing strategy needed to effect communication under certain discourse circumstances. (This emphasis on discourse constraints as organizing grammar is also an important feature of Karmiloff-Smith [1981], whose subjects produced quite different grammatical features in connected discourse as contrasted with their single sentence utterances.) An analysis of the "progressively more-syntacticized registers" posited by Givón could provide us with a more principled set of guidelines for what we might

expect in second language learners, especially at intermediate and advanced levels.

For several reasons, however, the role of discourse in organizing the acquisition of grammatical processing should not be over-emphasized. First, discourse constraints extend but do not override constraints on grammaticality at clause and sentence levels. A particular discourse context may dictate which of two possible grammatical structures is appropriate, but it will not normally allow the choice of a structure which is ungrammatical at the clause or sentence level. Second, too heavy a focus on discourse may lead us to overlook learner problems with basic grammatical relations and other primary building blocks of simple clauses. Our textbooks routinely assume that these pose no particular problem because it is taken for granted that even learners who have not been exposed to advanced grammatical terminology are familiar with the concepts traditionally denoted by terms like subject, direct object, and indirect object. Garrett's research (1982) indicates that this assumption is groundless, since even good language learners often have conceptually mistaken interlanguage rules for these relations. Traditional explanations of grammatical relations (to whatever extent they are given at all) usually consist of a confused conflation of semantic and syntactic properties, failing to indicate (either explicitly or through coherent presentation of examples) that there is a difference between semantic and syntactic roles, and that languages vary considerably as to the assignment of semantic roles to syntactic case. English, for example, allows both beneficiary and experiencer to be represented as grammatical subject, but most of the commonly taught languages do not. Even the processing of such basic notions as time and space is by no means transparent to the naive learner.

Summary

The cumulative import of the theory and the research cited here suggests that our theoretical and practical problems with the role of grammar in the development of communicative ability might be effectively addressed if we adopted a perspective establishing that:

(1) processing or performance, rather than competence or knowledge, is the central construct in acquisition;
(2) interlanguage is a set of processing rules rather than knowledge rules;
(3) the processing of grammatical form is not necessarily different from that of lexical items;
(4) grammar is not per se separate from meaning and should not be seen as a formal entity to be taught or

assessed separately from the communication of semantic, sociolinguistic, discourse, and pragmatic meaning;

(5) both linguistic competence and communicative competence are knowledge structures, not abilities.

This discussion of the processing perspective on grammar should suggest serious grounds for concern with the way grammar is represented in the Proficiency Movement. As discussed earlier, its inadequately principled basis for assessing grammatical competence within communicative competence is representative of the inadequacy of the field as a whole in this regard. There is good reason for the absence of grammatical specifications for the realization of communicative functions in the generic guidelines; but in the language-specific ratings, the description of accuracy poses a real dilemma. Not to include grammar specifications is to ignore the form end of the form-meaning connection, but to include them is to base ratings on the assessment of product rather than process, and, moreover, tempts untrained teachers to use the OPI as a discrete-point elicitation task.

A more serious problem is that the levels are inherently sequential and implicational, while the acquisition of processing rules is not. There is no way around the achievement-test quality of the grammatical expectations at the lower proficiency levels, and no way around the circularity of the level descriptions with respect to the learner's supposedly syllabus-independent ability at any level. These points have been made before in relation to the assessment of other competencies.

We must consider what we can actually discover about a learner's ability to create functionally appropriate and appropriately accurate form-meaning connections on the basis of a TL interview. As to the meanings involved, since we are dealing with adult learners, we cannot assume any predictable progression in the complexity of the meanings they will want to and try to convey; they are cognitively quite mature if not yet fully skilled (even in the NL) at hypothesizing about abstractions. That is, we cannot assume that a second-year student is attempting to convey only meanings for which the grammatical form is taught in the second year. As to the formal features, I have argued that the appearance or non-appearance of particular forms in their obligatory contexts (linguistic or communicative) cannot be assumed to give us useful information about what the learner actually means by the form. As Valdman (1980) and Garrett (1986) pointed out, there is no psycholinguistic basis for the rating descriptions.

It must be concluded that the OPI is inadequate to test oral proficiency and that the use of the *Guidelines* to shape syllabus design and teaching practice is unjustified. I do believe that the OPI is a usable

oral elicitation device and that the data it yields could be used together with other data to contribute to a teacher's sense of a student's ability to perform in one kind of discourse situation. There is no question that the training sessions and workshops familiarizing teachers with the notion of oral proficiency and the multiplicity of factors which go into it have given teachers a more realistic understanding of the goal of communicative language ability and have given them techniques–to be adapted to their own students' situation–for assessing some aspects of what they are trying to teach. Beyond that, however, the uses to which the test and the *Guidelines* are widely being put are unwarranted.

Conclusion

It is part of the protocol to end an article like this one with the call for more research, but it is pointless and potentially divisive to call for research without addressing the questions of what kind of research is most urgently needed, how it can be carried out, and by whom, and what we can expect to do with the results. This article concludes, therefore, by positing some desirable features of research and theory development with respect to the role of grammar.

(1) Above all, we need to understand second language acquisition in our own terms, as it takes place in the adult foreign language classroom, rather than extrapolating from research on language acquisition under altogether different circumstances. (It goes without saying that our theory will have to be consonant with the language acquisition theory developed for other domains).

(2) We must clearly understand the distinction between *research in second language acquisition in the classroom* and *research in foreign language education* and insist on the priority of the former. That is, we must have some idea of what learners are actually doing with language before we can study the effect of manipulating pedagogical variables.

(3) The suggestion that explanation or error correction based on the concept of processing instead of on linguistic rules will have more effect is a testable hypothesis. Most of the arguments against explicit grammar rest on the illogical conclusion that, because the application of explicit traditional linguistically based grammar has a limited effect on learners' communicative ability, we should not try to teach any kind of grammar

or attempt any kind of error correction. Perhaps there are certain stages of processing (or certain structures, or certain kinds of meaning) which will turn out to be resistant to pedagogical manipulation–if so, it would be extremely interesting to know what they are and why–but that certainly cannot be assumed on the basis of research to date.

(4) Research must focus on the individual learner or on particular characteristics by which data from individuals can validly be grouped as well as on data aggregated across groups. There are too many individual characteristics, such as cognitive styles, affective traits, and language learning backgrounds, whose influence on processing is as yet unknown. Generalizations about language acquisition by groups must be complemented and informed by more information on individuals.

(5) Once we have the theoretical framework and methodological sophistication to investigate processing rules, we can study their differential development under a variety of pedagogical treatment. We need to study three approaches, all in classrooms organized from a communicative perspective: (a) where grammar is not explicitly taught in any way, (b) where grammar is taught as it is now in communicative classrooms whose teachers do believe a little is worthwhile, and (c) where the grammar presentations are explicitly organized from the processing perspective.

There are many reasons, too well-known to require listing here, why the undertaking of classroom second language acquisition research has been so difficult for all sectors of the field. One of them is the extraordinarily heavy time commitment demanded of language teachers at every level. With the rapid increase of teacher and student access to microcomputers, however, the time required to do research can be dramatically reduced. The potential of computer-assisted instruction in foreign language education is well-recognized, though as yet poorly realized. (Garrett, 1987, discusses the possibilities for realizing the concept of processing grammar in computer-assisted language learning materials.) The enormous potential for computer-assisted basic research, however, is as yet hardly dreamt of–not studies designed to evaluate the efficacy of computer-assisted instruction, but second language processing research which uses the computer to elicit, collect and, analyze learner language data. The technology has for some years already been capable of supporting processing research on complex

written data (cf. Garrett, 1982), and computer-controlled audio and video technologies now allow, in addition, aural and visual manipulation of both stimulus and response. It is the conceptual design of the research, not the implementation, which is difficult. All it takes now is for teachers to recognize the potential and learn how to use it.

The research agenda outlined here is considerably broader than an investigation of the role of grammar in the development of communicative ability, but the organizing principle of processing allows us to see that grammar in its widest sense includes all the ways in which words are put together to communicate meaning. The most serious challenge facing the Proficiency Movement is the need to integrate into the effort to improve teaching and testing an ongoing program of research and theory development dedicated to a holistic concept of language acquisition which will finally include a principled concept of the role of grammar.

Notes

[1] Canale and Swain (1980) give a careful overview of various theories of communicative competence which contribute to our present understanding of it, and they distinguish between those theorists who consider communicative competence to include grammatical competence along with sociolinguistic, pragmatic, and discourse competencies and those who see the communicative as separate from, complementary to, the grammatical. As will become clear later, I adhere to the former.

[2] In the discussion of several other papers presented at the Symposium, the meaning attached by the PM to the term "accuracy" was the focus of considerable debate. Omaggio insists (personal communication) that it is intended to include accuracy of vocabulary and pronunciation as well as of grammar, and appropriateness in terms of sociolinguistic and discourse competence as well. However, in foreign language education generally as well as in linguistics, "accuracy" usually refers to grammar, and "appropriateness" refers to sociolinguistic, pragmatic, and discourse factors, which is how I use the terms in this paper.

[3] This metaphor itself implies the second view of communicative competence discussed in Note 2 above, the notion that communicative rules are separate from and added onto grammatical ones. This view is explicitly rejected by Hymes (personal communication) and others, myself included.

[4] In comprehension, the processing rules operate to decode utterances into meaning, although probably not in a crudely reversed fashion.

[5] See Garrett (1982, 1986) for a fuller development of this argument.

[6] Another important variable in the learning situation, the different challenge posed to native speakers of English by the learning of highly inflected vs. less highly inflected languages, of typologically similar vs. dissimilar languages, etc., has as yet received almost no explicit research attention in foreign language education. A very high proportion of serious work in second language acquisition theory and empirical research in interlanguage has been done on English, and of the work on the acquisition of other languages, a high proportion has been done on French and Spanish, languages which are both closely related and relatively less inflected. It should not be assumed that second language acquisition theory which has been developed on the basis of this limited research (still less any general notion about how much explicit attention should be paid to grammar) is generalizable to all classroom language learning.

References

Corder, S. P. (1977). Simple codes' and the source of the second language learner's initial heuristic hypothesis. *Studies in Second Language Acquisition, 1*.

Garrett, C. [same author as Garrett, N.] (1982). *In search of interlanguage: A study of second-language acquisition of German syntax*. Unpublished doctoral dissertation. University of Illinois, Urbana-Champaign.

Garrett, N. (1986). The problem with grammar: What kind can the language learner use? *Modern Language Journal, 70*; 133-148 (esp. Note 23).

Garrett, N. (1987). Psycholinguistic grammar and CALL [Computer-Assisted Language Learning]. In W. F. Smith (Ed.), *Modern media in foreign language education: Theory and implementation*. Lincolnwood, IL: National Textbook.

Givón, T. (1979). From discourse to syntax: Grammar as a processing strategy. In T. Givón (Ed.), *Syntax and semantics 12: Discourse and syntax*. New York: Academic Press.

Karmiloff-Smith, A. (1981). Language as a formal problem space for children. In W. Deutsch (Ed.), *The child's construction of language*. London: Academic Press.

Kelly, L. (1969). *25 centuries of language teaching*. Rowley, MA: Newbury House.

Slobin, D. T. (1977). Language change in childhood and in history. In T. Macnamara (Ed.), *Language learning and thought*,. New York: Academic Press.

Valdman, A. (1980). Communicative ability and syllabus design for global foreign language courses. *Studies in Second Language Acquisition, 3*, 81-95.

Author

NINA GARRETT, Research Associate, Language Learning Laboratory, University of Illinois at Urbana-Champaign, G70 FLB, 707 South Mathews, Urbana, IL 61801. Specializations: psycholinguistics of classroom language acquisition, computer-assisted language learning, and computer-assisted language acquisition research.

Applied Language Learning
Spring 1989, Vol. 1, No. 1, pp. 33-51

The Problem of the Target Model in Proficiency-Oriented Foreign Language Instruction

Albert Valdman
Indiana University

The article deals with the construct of educated native speaker (ENS) speech, which constitutes the implicit target norm in the ACTFL Guidelines and the explicit one for the ILR version. It is argued that in classroom FL instruction such a target is too restrictive, and as an alternative, a multitarget model of language variation is provided, according to which, speakers differ with respect to norms and the degree of influence exercised on them by shared norms in particular social contexts. It is pointed out that the speech of native speakers varies with regard to relative planning in response to another set of social and contextual factors. In addition to the well documented variation determined by the nature of linguistic tasks, FL learners' interlanguage is influenced by variable orientation toward the competing norms of the target language community. The article suggests that the special social context represented by the FL classroom warrants the establishment of special pedagogical norms. Sociolinguistically and psycholinguistically based principles for the elaboration of these norms are set down, and their application is illustrated in the case of French wh-questions, a notoriously variable feature of that language. Because it involves the presentation in pedagogical materials of sociolinguistically stigmatized features, the notion of pedagogical norm raises the issue of fossilization. It is noted that sociolinguistic deviance is unavoidable in early-level interlanguage; even speech targeted on ENSs may be deviant in certain social contexts. The article concludes with the suggestion that fossilization may be reduced by a better match between communicative demands and linguistic means available to learners: the

---

A preliminary version of this article appeared in A. Valdman, (Ed.), (1987), *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency*. Bloomington, IN: Committee for Research and Development in Language Instruction. Grateful acknowledgement is made of the research contract from the Defense Language Institute which provided partial support for the Symposium.

former are not lessened but rather the latter are fitted to the learners' limited processing capabilities.

Learning to communicate in a foreign language is a perilous adventure in which the voyager must navigate a straight course between the Scylla of grammatical inaccuracy and the Charybdis of functional ineffectiveness. The introduction of the notion of communicative competence has led second language teaching methodologists to give higher priority to semantic appropriateness and functional adequacy. For example, in defining the Threshold Level, Van Ek (1975) placed emphasis on getting meaning across in communicative interactions and downgraded the importance of linguistic well-formedness. In the ensuing years hasty and incautious adaptations of the notional-functional syllabus to the American foreign language teaching situation have led language teachers to accept grammatically inaccurate production on the part of learners provided that it got the message across. More recently, however, more balanced views have prevailed which recognize the inseparability of grammatical accuracy, semantic interpretability, functional adequacy, discursive coherence and cohesion, and sociolinguistic appropriateness.

The fundamental issue in developing tests of communicative ability and in devising guidelines for the imparting of communicative skills is not the central versus the peripheral nature of grammatical accuracy or the legitimacy of a grammatically-oriented syllabus, but rather the identification of target norms for grammatical well-formedness and the specification of minimum levels of linguistic competence (Canale & Swain, 1980).

The *ACTFL Proficiency Guidelines* (in both the provisional and the 1986 versions) have set speech typical of Educated Native Speakers (ENS) of the target language as the implicit target for foreign language learners. As the *Guidelines* recognize, ENS are capable of perceiving and producing other varieties of the language.[1] What does not appear clearly in the *Guidelines* is that, in addition to shifting to other norms, ENS show highly variable usage. What we call ENS speech is in fact an idealization from highly variable use. These target users, to differing degrees, control a repertoire of grammatical forms, discourse styles, and sociolinguistic rules which they apply variably, depending on the situational context. It is control of this repertoire, not of a particular norm, that should constitute the ultimate goal for foreign language learning.

In this article, I will show that the ENS target chosen by the *ACTFL Proficiency Guidelines* eliminates from consideration a range of native speech which may constitute more appropriate targets for certain learners, particularly at lower proficiency levels. I will then

develop a model of interlinguistic variation that legitimizes non-standard varieties of the target language, and I will argue that, given the special nature of the foreign language teaching situation, the traditional standard norm should be replaced by dynamic pedagogical norms based not only on ENS usage but also on other norms of the target language speech community. I will conclude by showing that this notion of pedagogical norm makes possible at lower proficiency levels a higher degree of grammatical accuracy by providing a better match between the target and learners' capabilities.

The ENS as Target Model

In the *Guidelines* framework, the ENS constitutes the ultimate target for second-language learning. According to the *Guidelines* the chief difference between ENS and other speakers who are not such worthy models for foreign learners resides in ENS' more extensive repertoire. ENS have the capability of perceiving and producing features that characterize other varieties of the language. For example, the characterization of speaking ability at the superior level specifies that learners are "partially familiar with regional or other dialectical (sic) variants." In the area of comprehension, upper levels of proficiency are specified in terms of the ability to "understand all forms and styles of speech pertinent to personal, social, and professional needs tailored to different audiences." It is pointed out, however, that even at the distinguished level second-language learners "may have difficulty with some dialect and slang." Implicit in the sociolinguistic model that underlies the *Guidelines* is the notion that ENS are able to shift in and out of their habitual speech styles to produce speech varieties associated with social groups located below them in the social scale or in geographical areas outside of the prestige centers. Presumably, members of other social groups or residents of other geographical areas of target language (TL) countries do not possess this stylistic manoeuver or the ability to understand or produce the speech of other groups in the overall speech community.

The notion of ENS as more skillful and deft users of the language also appears in the only explicit discussion of the target norm within the *Guidelines* framework (Lowe, 1985). For Lowe, speakers of a language operate at the S-3 level in most of their daily communicative interactions; only occasionally do they move into higher levels. However, Lowe also correctly links ENS status to formal educational training and, especially, social class. He points out that the concept of ENS applies more readily to *stratified* societies such as those of France, Germany, and Great Britain, than those, such as that of the United States, that have undergone social leveling.

The belief that educated middle class speakers possess a richer speech repertoire than members of other social groups has affected the design of sociolinguistic research. Not surprisingly, in view of the bias in favor of ENS, empirical studies often confirm the lack of variability in the speech behavior of lower class speakers (Lindenfeld, 1972; Behnsted, 1973). Lindenfeld found that French speakers from the lowest of the two social classes she had identified showed the same level of syntactic complexity in formal and informal style. But she induced so-called formal speech by asking informants to imagine they were presenting a lecture before an auditorium audience of a hundred persons. This is hardly a situation that is within the realm of experience of working class speakers, and it provides little information about syntactic features that such speakers actually produce in genuine formal interactions characteristic of their social group. Behnsted designed his study of variation in French interrogative constructions on the assumption that middle class speakers showed different syntactic structures in formal versus informal style whereas their working class compatriots operated with a single style, see Table 1.

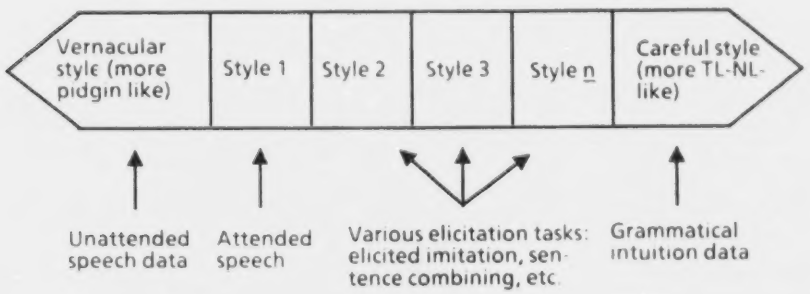
TABLE 1
Relative frequency of the distribution of interrogative constructions in a representative corpus of spoken French (Behnsted, 1973)

Type of construction			Working class	Middle class Informal	Formal
FRONTING	Où tu vas?	Où Jean va?	36%	46	10
PRONOMINALIZATION	Tu vas où?	Jean va où?	12	33	25
EST-CE QUE					
Est-ce que	Où est-ce que tu vas? Où est-ce que Jean va?		8	12	3
Que	Où qu' tu vas?		26	--	--
C'est que	Où c'est qu' tu vas?		3	4	--
Que c'est que	Où que c'est que tu vas?		6	--	--
INVERSION					
Simple	Où vas-tu? Où sont passes mes cigares?		9	3 2	47 12
Copy of	Où Jean va-t-il?		--	--	3
PRO			(N-587)	N-446)	(N-436)

Two central elements of Labov's variationist model (1966) bear on issues in second language learning related to the Proficiency Movement. First, speakers' realizations of a variable are arranged along a continuum reflecting degree of attention paid to speech. Second, the vernacular pole, reflecting the least attention paid to speech and showing least variation, represents speakers' most regular and systematic linguistic patterns. The few available studies of variation in second language learners' interlinguistic production have adopted these two central hypotheses of Labovian sociolinguistics. L. Dickerson (1975), W. Dickerson (1976), and L. and W. Dickerson (1977) have examined the development of phonological interlinguistic features over time and in different situations of speech production. They show that, in drawing closer to the TL model, learners follow a course analogous to that of phonological change and to the synchronic variation found in NS speech: they produce a variety of replicas for given TL features, certain replicas predominating in a particular environment. Over time, and differentially across speech styles, replicas approximating the TL feature will spread progressively to new environments, just as in a natural language standard variants spread differentially across various environments, word classes, social groups, or geographical areas.

On the basis of experimental studies focusing on morpho-syntactic as well as phonological features, E. Tarone (1979, 1983) hypothesizes that learners' interlanguage (IL) constitutes a continuum analogous to the Labovian vernacular-standard one. The relative position of an interlinguistic variant on the continuum is a function of the attention paid to linguistic form, see Figure 1. The left-most pole of the IL continuum corresponds to the vernacular on the natural language continuum. Vernacular interlinguistic features are produced when attention to linguistic form is lessened, and these features exhibit greater regularity. On the contrary, the right-most pole, corresponding to the standard end of the native vernacular- standard

FIGURE 1
The Interlanguage (IL) Continuum(from Tarone, E. 1983, p. 152)



continuum, is likely to contain TL and L1 features and shows the least internal consistency. Like Corder (1977) and Widdowson (1976), Tarone accounts for the internal consistency of the IL "vernacular" in terms of universal language processing strategies and substantive universals appearing under conditions of reduced input, such as those that trigger pidginization.

A Multi-Target Model of Language Variation

Multi-Target Orientation in Vernacular Speech

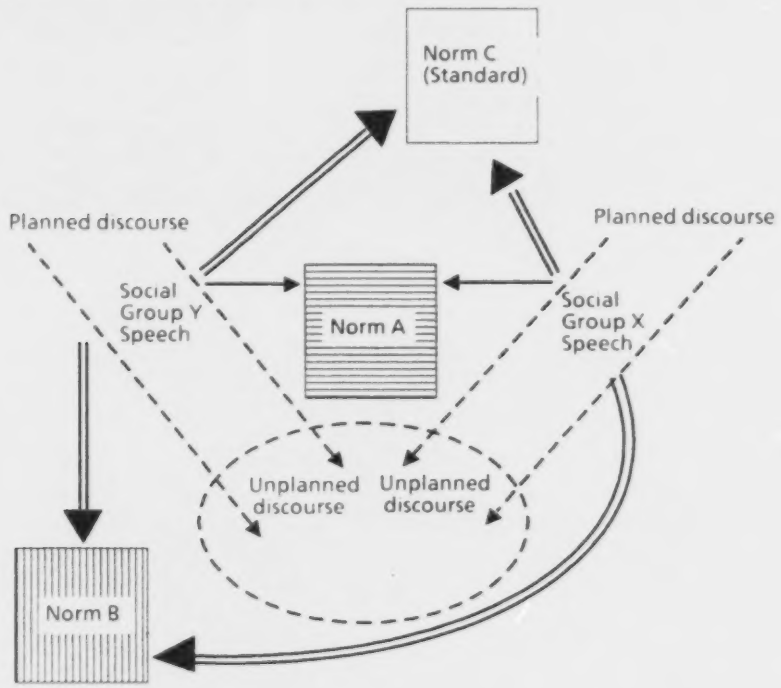
The early Labovian model adopted by Tarone needs to be enriched so that it takes into account linguistic accommodation and the existence of a variety of norms within any complex linguistic community. In his study of urban Black English, Labov (1972, p.249) accounts for the persistence of stigmatized forms in the vernacular speech of Black youths by invoking a set of covert norms that attribute positive values to the vernacular. He also remarks that verbal behavior determined by these norms is unlikely to appear in the formal situations that characterize sociolinguistic interviews. The following case illustrates vividly that non-standard speakers may be influenced by a variety of norms, including the more prestigious standard form.

In casual observations of linguistic variation in a small village in the Touraine region of France, Pierre Léon (1973) noticed that a group of male speakers from farmer and worker backgrounds shifted away from local vernacular speech in two directions. For example, they replaced the local apical tap *r* either by the velar resonant of standard French or by the pharyngeal resonant characteristic of lower and working class Parisians. The latter feature coincided with macho behavior and took place typically at meetings in a café run by a retired railway worker from Paris, a former militant Communist union organizer. This social movement away from the vernacular proceeded in two different directions, one oriented toward ENS speech (the standard norm), and the other toward what may be termed Parisian proletarian speech. Shifts toward one or the other of these two orientations cannot be accounted for exclusively in terms of relative attention to linguistic form. The triggering factors are complex and must be sought in the nature of the social relationships among participants in the speech event and in the situational context.

As an alternative to the unidirectional vernacular-standard continuum model, I propose a multi-target model, see Figure 2. I posit that all social groups in a speech community have available a range of

speech styles and that these reflect the influence of several norms. In Figure 2, the square boxes labeled Norm A, Norm B, and Norm C, respectively, represent alternative norms available to the various social groups of a speech community. For the sake of simplicity, the model posits only two social groups, *X* and *Y*. The width of the arrows reflects the relative power of attraction exercised on a particular social group by a given norm. Shifts in norm orientation are determined by the social identity speakers wish to signify in the course of given speech events. In daily communicative interactions with members of their own social groups, speakers make use of their vernacular, which is targeted on the norm set by members of the in-group. In complex societies, one of the norms available in the overall speech community is endowed with prestige and/or political power. This norm also undergoes a gradual progress of uniformization and gains the status of standard. The norm also becomes associated with the dominant social group, whose vernacular shows less deviation from it than that of the other social groups; in Figure 2, Group *X* represents the dominant group of the

FIGURE 2
Multi-target model of language variation



society whose speech networks are modelled. In the case cited by Pierre Léon (1973), the vernacular speech of local farmers is oriented toward the local rural Touraine norm, but they will shift toward the Parisian proletarian norm or Standard French, depending on the *acts of social identity* in which they engage in a particular situation (Le Page & Tabouret-Keller, 1984).

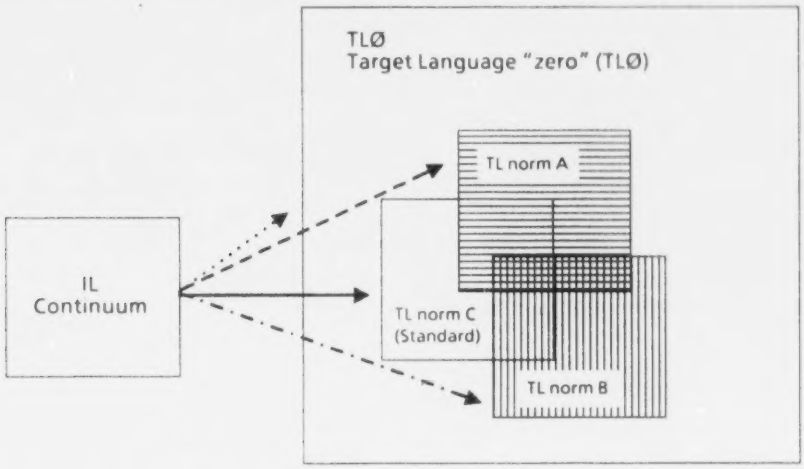
In addition to reflecting orientation toward a variety of norms, a speaker's verbal production varies between planned and unplanned discourse. Unplanned discourse, which characterizes conversational interaction, shows the predominance of reliance on the situational context and on lexicon and, consequently, syntactic reduction (Ochs, 1979). Planning involves attention to speech form, so that the distinction I am making between planned and unplanned discourse corresponds somewhat to that made between attended and non-attended speech in Tarone's model (1983), see Figure 1 above. The chief difference between my model and hers lies in the independence as opposed to the convergence of two distinct continua: unplanned → planned discourse (non-attention → attention to speech form) on one axis and vernacular → external norm on another. By failing to recognize the independence of these two axes of speech variation some linguists have been led to assume that non-standard speech lacks syntactic organization (Stourdzé, 1969; Bonin, 1978). Also, as was the case in the Lindenfeld (1972) study, they did not provide for the possibility that non-middle class speakers could produce both planned and unplanned discourse. As would be expected, because it relies on shared information and is linked to the *hic et nunc* of a specific communicative situation, the unplanned discourses of different social groups converge. Conversely, we would expect to find the greatest linguistic differentiation between two social groups to be located in their planned discourse.

A Multi-Target Norm for IL Variation

Within the framework of the multi-target norm model I propose for TL variation, the IL continuum may be viewed as a vector which is oriented toward a particular TL norm by filtering input and controlling feedback, see Figure 3. The largest square represents the totality of attested TL variants. Following R. Chaudenson (1986) I call this system TLØ. This system encompasses features shared by all TL varieties as well as all variable features. The three smaller squares represent competing norms, one of which is the standard. Line C represents an interlinguistic continuum heavily controlled by standard norm input, whereas lines A and B represent interlinguistic systems oriented toward other norms. In naturalistic second language learning,

the types of communicative situations encountered by the learner will determine in large part the orientation of the IL continuum. In guided learning, the orientation may be controlled to a certain extent by the selection and sequencing of linguistic features, by classroom procedures, and more particularly by the elaboration of pedagogical norms.

FIGURE 3
Possible orientations of IL continuum toward various TL norms



The Elaboration of Pedagogical Norms

The Concept of Pedagogical Norm

There is a growing recognition on the part of second language acquisition researchers that the foreign language classroom constitutes a special environment for second language development (Ellis, 1984; Faerch & Kasper, 1985; Edmonson, 1985; Kramsch, 1985). The varieties of the TL that are appropriate for classroom instructional and communicative activities are determined by the attitudes of the TL community, by those of the learners themselves, and by factors inherent in these activities (Soulé-Susbielles, 1984).

Whereas there exists a considerable body of research on native speakers' reactions toward learners' interlinguistic features (see d'Anglejan, 1983 for a review of the literature), socio-psychologically oriented studies of native speaker attitudes have not yet been extended to the area of language variation and the definition of a

sociolinguistically acceptable learner repertoire for the TL. From the only available exploratory study (Swacker, 1976), it would appear that native speakers tend to downgrade foreigner speech that diverges markedly from the standard norm. Swacker observed that native speakers of East Texas English reacted negatively to a reading sample from an Arabic accented speaker whose speech contained local pronunciation and grammatical features. A higher ranking was assigned to another Arabic accented speaker whose English was free of local linguistic features. Swacker concluded that certain regional markers may be acceptable when found in the speech of native speakers but are offensive in the mouths of foreigners. Swacker fails to specify, however, whether the local features noted also symbolize low social status within the TL community.

Without a substantial body of empirical research, we must resort to extrapolation from other areas of interlinguistic research to seek principles for the determination of TL norms suitable for foreign learners. The reaction of Swacker's native judges may be accounted for in terms of a model of status and solidarity (Ryan, 1983). Foreign accents and other interlinguistic features are highly salient markers of membership in outgroups. Their presence in the speech of a foreigner would preclude features symbolizing ingroup membership, such as regionally or socially marked variants. Appeal to the notion of linguistic capital (Bourdieu, 1977) also supports the selection of a highly standardized TL norm for foreign learners. The learning of the TL may be viewed as an economic investment whose value would be depreciated if the variety mastered contained stigmatized features. The higher the social status conferred by the learned speech forms and the communicative strategies, the more profitable the investment.

These various considerations suggest that the language used in the classroom should differ from the naturalistic varieties found in communicative interactions among native speakers. There are differences between what may be termed a *pedagogical norm* and the various norms that determine which native speaker usage will be particularly marked at early stages of instruction when learners' speech shows many deviant features, some attributable to transfer from the native tongue, others reflecting extensive restructuring of TL input, and still others residing in the communicative strategies specific to the classroom context. Thus, a central task in organizing classroom FL instruction is the elaboration of pedagogical norms, particularly at early stages.

A pedagogical norm, like all norms, is an abstraction. Its distinguishing feature resides in the fact that it is an artificial construct reflecting the special conditions of classroom foreign language learning (Valdman, 1976, 1983). Four sets of principles guide the elaboration of pedagogical norms. First, they should reflect the actual behavior of TL

speakers in authentic communicative situations. Second, they should conform to native speakers' idealized views of their own linguistic behavior. Third, they should match both TL speakers' views about what linguistic behavior is appropriate for educated foreigners and the perceptions of learners themselves. Fourth, they should take into account processing and learning factors, such as learners' interlanguage features, since these point the way to what is most easily learnable.

The elaboration of a suitable pedagogical norm requires first of all a clear determination of learning objectives and a careful review of the social and broad contextual factors that are likely to shape the particular classroom environment. From a more narrowly linguistic and psycholinguistic perspective, it requires linguistic descriptions of a level of delicacy not attainable by the application of current linguistic models. Not only must applied linguists be supplied with information on the semantic, pragmatic, and sociolinguistic value of specific target language features but they must also be given insight into the relative optimality of the available structural options. I am using the term *optimality* to refer to ease of acquisition and centrality within the overall structure of the TL. This notion no doubt corresponds in part with markedness in the sense that optimal features would be expected to be unmarked from the point of view of language universals.

Illustration of a Pedagogical Norm

In this section, I will summarize the application of a pedagogical norm to a notoriously variable aspect of French syntax–*wh*-questions–which I have treated elsewhere (Valdman, 1975, 1976, 1983, 1986).

Traditionally, two *wh*-constructions are selected for the teaching of French as a foreign language: *est-ce que* and inversion. However, as Behnsted's sociolinguistically oriented study shows (see Table 1), these constructions are relatively rare in vernacular speech. Instead, native speakers use two variants often characterized as *ungrammatical*, but which are syntactically less complexified: pronominalization and fronting. It is noteworthy that ENS (Behnsted's middle class informants) overestimated their use of the most valorized construction, inversion, and, inversely, underestimated the frequency of the two devalorized variants.

There are no studies of NS attitudes regarding the relative appropriateness of the *wh*-question variants for foreign learners. Extrapolating from Swacker's pilot study (1976) and referring to the criteria of status and solidarity (Ryan, 1983), one may assume that French ENS would expect educated native speakers to adhere closely to their own subjective norm and favor the two constructions associated with formal styles. From a psycholinguistic perspective, data from two

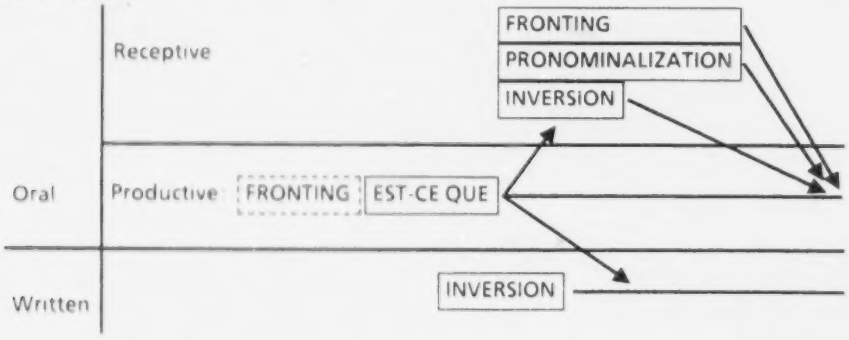
different lines of inquiry suggest that the two most frequent *wh*-question types–pronominalization and fronting–are the most easily learnable. First, the latter construction is universally found in European-based creoles where vernacular varieties of the lexifier language show different syntactic arrangements from the creole. For example, compare *Where have you been?* and Guyanese Creole *Wisaid yu bin de?* (Bickerton, 1981, p.70). Second, a study of the acquisition of French *wh*-questions by American English learners (Valdman, 1975, 1976) revealed a high proportion of fronting. That structure occurred in utterances produced within the framework of a test of oral ability administered to beginners after approximately 75 hours of instruction. The presence of the structure represents a case of creative construction (Dulay, Burt, & Krashen, 1982), since the learners had never been exposed to it; the construction was absent from the course materials and from the speech of instructors, all of whom used inversion or *est-ce que* in both drill and communicative activities. The presence of the feature cannot be explained by transfer from the L1 either, since English *wh*- questions may be described, at the surface level, either as inversion or as the adjunction of a semantically empty marker *do* matching *est-ce que* with respect to position:

Where	*does*	John	work?
↕	↕	↕	↕
Où	*est-ce que*	Jean	travaille?

The application of the four criteria proposed for the elaboration of pedagogical norms suggests the following dynamic norm for the introduction of *wh*-questions. The norm is dynamic in the sense that the learner's linguistic behavior would be expected to draw progressively closer to the selected target. The first three criteria for a pedagogical norm converge to identify inversion as the most highly valued structure from a sociolinguistic point of view. No doubt native speakers would deem it most appropriate for foreigners who have acquired the language by formal instruction. Bourdieu's view of language as capital would also lead learners to assign it the highest surrender or market value. On the other hand, it is highly marked, both from the syntactic and the sociolinguistic point of view. Its close link with formal discourse and written texts render it inappropriate as the target feature in elementary and intermediate courses which have as one of their main objectives the acquisition of a certain degree of communicative ability. The numerous syntactic constraints that apply

to inversion (Valdman, 1975) render it difficult to learn and to use with accuracy. Its sociolinguistically neutral status and its relative syntactic simplicity designate *est-ce que* as the first-level target feature. However, since fronting represents a case of creative construction, it is suitable for transitional use in early stages of instruction, see Figure 4. Inversion would be introduced at Stage 2 in lexicalized frozen forms in cases where *est-ce que* is judged too awkward; for example compare: *Comment allez-vous? 'How are you?' (Comment est-ce que vous allez?)*, *Où vont vos parents? 'Where are your parents going?' (Où est-ce que vos parents vont?)*. At that stage, inversion would also serve as the canonical form for written expression. At more advanced levels (Stage 3), pronominalization and fronting would be identified as more appropriate for informal situations. As is the case for inversion, the use of pronominalization and fronting in contrived materials should reflect their distribution in authentic, unedited speech samples so that learners may be led to induce their pragmatic and sociolinguistic values. Figure 4 represents a proposed pedagogical progression indicating the order of introduction of the four variants in speech and writing, in which a distinction is made between receptive and productive phases.

FIGURE 4
A proposed pedagogical sequence for the presentation of French interrogative constructions



	Stage 1 (Preliminary)	Stage 2	Stage 3	Stage 4
Frozen utterances	Quelle heure est-il? Comment vous appelez-vous? Quel temps fait-il?			

The concept of pedagogical norm makes it possible to orient the IL continuum toward the ultimate TL norm selected. The TL variant selected for initial presentation is the one that learners can most easily process. Should it be a stigmatized feature, it is progressively replaced by a sociolinguistically neutral or valorized variant. In the specific case of French *wh*-questions, fronting would be selected as the first target feature. It would be replaced by *est-ce que* as soon as the learners have acquired control of *wh*-words. Inversion would be introduced in written discourse and, in the final stage, all four constructions would be presented for recognition and active control in appropriate situational contexts. Any particular group of learners would need to demonstrate the ability to fit the *wh*-question variant construction to the intended audience and the specific situational context. For example, if the instructional objective were to impart a level of linguistic proficiency sufficient to enable Foreign Service officers to perform Level 4 tasks, these learners would need to attain Stage 4 in Figure 4. On the other hand, the IL continuum would need to be set differently in the case of the average high school or college level student, for whom Stage 2 or early Stage 3 would be sufficient to carry out Level 2 tasks, for example, to engage in all types of casual conversations.

The Issue of Fossilization

A major issue in the Proficiency Movement is that of fossilization. Higgs and Clifford (1982) report that many trainees in governmental agencies' language programs who begin instruction with a relatively high level of proficiency (S-1+) fall short of attaining the critical S-3 level because lexical and grammatical features imperfectly mastered during previous study or language use become fossilized. To remedy this situation, Higgs and Clifford advocate more stringent attention to linguistic accuracy at initial levels of second language development. From that perspective, one could object to the application of pedagogical norms in which learner performance is oriented toward features characteristic of unplanned and non-standard vernacular models, as the result would be fossilization of these features. Learners would fail to attain the control of variants characteristic of ENS, whose planned speech is generally targeted on the standard norm. This type of fossilization would constitute a serious problem in the training of diplomats whose job description required negotiating with ENS counterparts.

I would respond to this type of objection by underscoring the distinction between IL features that fall outside of the construct I have labeled TLO (see Figure 3) and those within it. The former features are indeed linguistic inaccuracies to which the standard notion of fossilization applies. Sociolinguistically stigmatized features, which

abound in the speech of NS, may at worst be sociolinguistically or discursively inappropriate. But sociolinguistic and discursive inappropriateness also result when variants that characterize planned, formal discourse occur in situations that require unplanned discourse; such is the case, for example, for French inversion *wn*-questions. It is only the traditional enshrining of ENS planned discourse which leads to labeling features characteristic of non-standard speech and unplanned discourse (often termed slip-shod) as grammatically inaccurate. In addition, discussions of fossilization in IL fail to distinguish genuine instances of this phenomenon from the syntactic reduction linked to reliance on the situational context that accompanies shifting down from planned to unplanned discourse. In the context of the Oral Proficiency Interview (OPI), one wonders whether some cases of fossilization are not artifacts of the testing situation. At lower proficiency levels, the OPI can at best simulate naturalistic conversational interaction. Genuine conversational interaction is more likely to elicit unplanned discourse. Indeed, planned discourse would be artificial in that type of situation.

If anything, reliance on the situational context and on lexicon and accompanying reduction of syntax are pragmatically compatible and sociolinguistically appropriate in a test portrayed as an approximation of genuine conversational interaction. It may be that learners afflicted with the so-called terminal 2/2+ syndrome might perform with greater grammatical accuracy in test situations that more closely model situations requiring the production of planned discourse.

Conclusion

Higgs and Clifford (1982, p.78) propose the "output hypothesis," analogous to Krashen's optimal input hypothesis (1982):

> When the communicative demands made upon the students are too far beyond their current competence, they are forced to adopt or invent communication strategies that lead to fossilization and ultimately prove self-defeating. For this reason, it is important to match the communicative task to the students' performance level.

I would agree that such mismatches between communicative and semantic intent and linguistic form must be avoided. However, the converse to the "output hypothesis" is equally dysfunctional:

> When second language learners are forced to use grammatical features and to make semantic differentiations that are beyond their processing capabilities, they resort to structural reduction. This leads to the production of

features, characteristic of unplanned discourse that appear structurally deviant from the perspective of ENS planned discourse.

It is the responsibility of syllabus designers, developers of teaching materials, and classroom instructors alike to avoid both types of mismatches between language use tasks and linguistic means available to learners. The notion of pedagogical norm I have developed in this paper promises to lead to this mutual adjustment between language form and language use.

Schulz (1986) voices serious reservations about the appropriateness of oral proficiency in a foreign language as a central goal in general education or liberal arts training. She argues that, in addition to the fact that it can only be attained in extensive sequences of study not usually found in U.S. schools and universities, speaking ability is the language skill most vulnerable to erosion or loss. On the basis of these and other considerations, I have proposed (Valdman, 1982) that, in addition to broad educational objectives, such as acquainting students with the multifarious links between language and mind, culture, and society, fundamental (as versus special purpose) foreign language instruction should more modestly and realistically aim at imparting minimal communicative capacity (MCC) rather than full-fledged communicative proficiency.

Within the framework of this type of fundamental foreign language instruction, the elaboration of pedagogical norms is all the more crucial. Clearly, MCC cannot be attained in a classroom context by the average learner if ENS planned discourse remains the target. That variety of the target language shows the highest level of syntactic and semantic complexification. At the level of syntax, many minor rules concerning subcategorization, government, and agreement must be acquired; and since speech will often be decontextualized, highly specific vocabulary items expressing fine nuances of meaning must be introduced. In contrast, unplanned discourse, which is highly context-embedded, features syntactic reduction and semantic simplification. A pedagogical norm comprises target language variants that involve the simplest syntactic machinery and that form the most regular pattern. Where social or geographical diversity exists, sociolinguistic factors will be abandoned in favor of psycholinguistic ones: those variants will be selected that are most easily processible by second language learners. It should be stressed that this approach does not involve the teaching of a pidginized version of the TL. The features that constitute the pedagogical norm do not have their origin in learners' interlinguistic approximations; they are drawn from variants occurring in a composite of actual native speech.

The approach I am advocating takes a neutral stand with respect to current debates about suitable classroom activities, whether communicative practice or modelling of a set of selected structural features, and about appropriate syllabus design, whether notional-functional or structural. For example, most speech acts may be expressed either by linguistic features characteristic of ENS or those of other less prestigious social groups; the same set of semantic notions may be overtly realized by the complex syntactic machinery typical of planned discourse or by the sentence fragments, formulaic combinations, and unmarked syntactic devices and morphological categories characteristic of unplanned discourse. The bottom line is that learners should be led to attain MCC as quickly as possible but in a manner that permits, in the long term, the development of language-specific marked linguistic devices that make decontextualized discourse possible. This objective will not be reached without a drastic overhaul of the linear structurally oriented syllabus that underlies most current teaching, including that which claims to be proficiency-oriented.

Note

[1]Lowe (1986) points out that in the parent ILR (Interagency Language Roundtable) the upper end of the scale is labeled *well-educated* native speaker (WENS). There does not seem to be any significant difference between the two concepts.

References

American Council on the Teaching of French Language. (1982). *ACTFL provisional proficiency guidelines*. Hastings-on-Hudson, NY: Author.

American Council on the Teaching of French Language. (1986). *ACTFL proficiency guidelines*. Hastings-on-Hudson, NY: Author.

Behnstedt, P. (1973). *Viens-tu, est-ce que tu viens? Formen und Strukturen des direkten Fragesatzes im Französischen*. Tubingen: Gunter Narr.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma Press.

Bonin, T. (1978). The role of colloquial French in communication and implications for language instruction. *Modern Language Journal, 62*, 90-102.

Bourdieu, P. (1977). L'economie des changements linguistiques. *Langue Français, 34*, 17-34.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Chaudenson, R. (1986). Norme, variation, créolisation. *AILA Review/Revue de l'AILA, 2*, 69-88.

Corder, S. P. (1977). Language continua and the Interlanguage Hypothesis. In S. P. Corder& E. Roulet (Eds.), *The notions of simplification, interlanguages and pidgins and their relation to second language pedagogy: Actes du 5ème Colloque de Linguistique Appliquée de Neuchâtel* (pp. 11-17). Geneva: Librairie Droz.

D'Anglejan, A. (1983). Introduction to special issue, native speaker reactions to learner speech. *Studies in Second Language Acquisition*, 5(2), vii-ix.

Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, *9*, 401-407.

Dickerson, L., & Dickerson, W. (1977). Interlanguage phonology: current research and future directions. In S. P. Corder & E. Roulet (Eds.), *The notions of simplification, interlanguages and pidgins and their relation to second language pedagogy: Actes du 5ème Colloque de Linguistique Appliquée de Neuchâtel* (pp. 18-30). Geneva: Librairie Droz.

Dickerson, W. (1976). The psycholinguistic unity of language learning and language change. *Language Learning*, *26*, 215-231.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. New York: Oxford University Press.

Edmonson, W. J. (1985). Discourse worlds in the classroom and foreign language learning. *Studies in Second Language Acquisition*, 7, 159-168.

Ellis, R. (1984). *Classroom second language development*. Oxford: Pergamon Press.

Faerch, C., & Kasper, G. (1985). Introduction to special issue, Foreign language learning under classroom conditions. *Studies in Second Language Acquisition*, *7*, 131-133.

Higgs, T. V., & Clifford, R. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum competence and the foreign language teacher* (pp.57-80). Skokie, IL: National Textbook.

Kramsch, C. J. (1985). Classroom interaction and discourse options. *Studies in Second Language Acquisition*, *7*, 169-183.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Léon, P. R. (1973). Réflections idiomatologique sur l'accent en tant que métaphore socio-linguistique. *French Review*, *46*, 783-789.

LePage, R., & Tabouret-Keller, A. (1984). *Acts of identity*. Cambridge: Cambridge University Press.

Lindenfeld, J. (1972). The social conditioning of syntactic variation in French. In J. Fishman (Ed.), *Advances in the sociology of language* (pp. 77-90). The Hague: Mouton.

Lowe, P. Jr. (1985). The ILR proficiency scale as a synthesizing research principle: The view from the mountain. In C. J. James (Ed.), *Foreign language proficiency in the classroom and beyond* (pp. 9-54). Lincolnwood, IL: National Textbook.

Lowe, P. Jr. (1986). Proficiency: Panacea, framework, process? A reply to Kramsch, Schulz, and, particularly, to Bachman and Savignon. *Modern Language Journal*, 70, 392-397.

Ochs, E. (1979). Planned and unplanned discourse. In T. Givon (Ed.), *Syntax and semantics: Vol. 12. Discourse and semantics*. New York: Academic Press.

Ryan, E. B. (1983). Social psychological mechanisms underlying native speaker evaluations of non-native speech. *Studies in Second Language Acquisition, 5*(2), 148-159.

Schulz, R. A. (1986). From achievement to proficiency through classroom instruction: Some caveats. *Modern Language Journal, 70*, 373-379.

Soulé-Susbielles, N. (1984). La question, un outil pédagogique dépassé? *Le Français dans le Monde, 183*, 26-34.

Stourdzé, C. (1969). Les niveaux de langue. *Le Français dans le Monde, 65*, 18-21.

Swacker, M. (1976). When (+ native) is (- favourable). *Lektos* [special issue], 135-254.

Tarone, E. (1979). Interlanguage and the chameleon. *Language Learning, 29*, 181-191.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 149-163.

Valdman, A. (1975). Error analysis and pedagogical ordering. In S. P. Corder & E. Roulet (Eds.), *Some implications of linguistic theory for applied linguistics* (pp. 105-126). Paris: Didier.

Valdman, A. (1976). Variation linguistique et norme pédagogique dans l'enseignement du français langue étrangère. *Bulletin de la Féderation Internationale des Professeurs de Français*, 12-13, 52-64.

Valdman, A. (1982). Toward a modified structural syllabus. *Studies in Second Language Acquisition, 5*, 34-51.

Valdman, A. (1983). Language variation and foreign language teaching: Issues and orientations. In L. MacMathuna & D. Singleton (Eds.), *Language across cultures* (pp. 171-184). Dublin: Irish Association for Applied Linguistics.

Valdman, A., Barnett, M., Holekamp, E., Laronde, M., Magnan, S., & Pons, C. (1986). *En Route - Introduction au français et au monde francophone*. New York: Macmillan.

Van Ek, J. A. (1975). *Systems development in adult language learning*. Strasbourg: The Council of Europe.

Widdowson, H. G. (1976). The significance of simplification. *Studies in Second Language Acquisition, 1*, 11-20.

Author

ALBERT VALDMAN, Rudy Professor of French & Italian and Linguistics, Indiana University, Bloomington, IN 47406. Specializations: second language acquisition, language contact, French, applied linguistics.

Applied Language Learning
Spring 1989, Vol. 1, No. 1, pp. 53-73

Testing Listening Comprehension in the Context of the ACTFL Proficiency Guidelines

Dan Douglas
Iowa State University, AMES

A discussion of issues related to the testing of listening comprehension in the context of the 1986 ACTFL Proficiency Guidelines (American Council on the Teaching of Foreign Languages, 1987) is presented in this article. After a review of current research on listening comprehension and development of listening tests, four issues are singled out for more detailed consideration: the meaning of context in listening testing, the concept of criterion-referenced tests, the notion of specific purposes in testing, and the use of technology in listening tests. A number of research issues are discussed.

This paper is a discussion of issues related to the testing of listening comprehension in the context of the 1986 *ACTFL Proficiency Guidelines*. Although a great deal of research has been conducted in recent years on the listening skill, and its central place in initial second language learning and influence on the development of the other language skills have become better understood, listening comprehension testing has been, and remains, somewhat the quiet corner in the testing field.

A number of researchers have provided insights into the complexity of the active listening skill: Richards (1983) considers *message* and *medium* factors to arrive at a list of 33 micro-skills required for conversational listening, and a list of 18 micro-skills relevant to academic listening; Chaudron and Richards (1986) show that macro-*markers* (those that highlight information in a lecture and its sequencing or importance) are more important to comprehension than are *micro-markers* (which indicate links between sentences in the lecture); Tyler (1987) showed that effective communicators use a

Studies in Second Language Acquisition, 10 (2), pp. 245-262. Copyright 1988. Reprinted by permission of the publisher.

number of prosodic, syntactic and lexical devices to orient their listeners to the relative importance and interrelationships of ideas in their discourse and that failure to use such devices led to a perception of disorganization; Glisan (1985) provides evidence for the importance of first- and second-language word order in the comprehension of key sentences in texts; Dunkel (1986) deals with the strategic processes of listening comprehension; and Brown (1986) reports on experimentation on the influence of context in listening comprehension. Powers (1986), in a report on a survey of the academic demands related to listening skills, outlines variables related to stimulus material, speaker, and context as they affect listening comprehension, and identifies nine specific skills as very important: identifying major themes or ideas, identifying relationships among major ideas, identifying the topic of a lecture, retaining information through notetaking, retrieving information from notes, inferring relationships between information, comprehending key vocabulary, following the spoken mode of lectures, and identifying supporting ideas and examples.

In short, one finds in current research a view that listening comprehension is a complex, active skill that is basic to and influential in foreign language learning. What evidence is there that this lively research interest has carried over into testing? What's going on currently in listening comprehension testing?

Current Work in Listening Comprehension Testing

In order to provide a brief overview of current work in listening test development and research, I employ the following categories:

1. *Listen and give the right answer tasks.* These are frequently multiple choice tasks, but need not be. They are employed with the assumption that testees will show that they can listen to a segment of speech and draw from it certain kinds of information. A problem with this assumption is that, frequently, no clear indication of purpose for the listening task is provided, beyond the situational one of the test itself.

2. *Reduced-redundancy tasks.* The assumption here is that testees will show they can comprehend a spoken text in spite of missing or masked features. One problem with the reduced redundancy task is that a normal response to garbled transmission is to try to tune in the radio station better or to dial a different one, unless there is a very important reason for wanting to listen to a particular one–such as a language test.

3. *Repetition/imitation tasks.* These tasks involve short term memory and are based on the assumption that testees' syntactic,

organizational competence will be reflected in their ability to repeat or imitate an oral input.

4. *Interaction tasks.* These tasks are based on studies of conversational interaction and assume that testees will produce or select appropriate responses in conversational situations or simulations.

5. *Media-transfer tasks.* The assumption in these tasks is that an ability to mediate between input media, say oral and pictorial, is reflective of a high-level cognitive/linguistic competence.

No rigorous theoretical implications are made by these categories; in fact, I would like to avoid, in this paper, getting mired in sorting out the many orientations, approaches, and bases that we currently find in the literature: for example, notional/functional, proficiency based, task based, content based, performance based, pragmatic and communicative. As Alan Davies (personal communication) pointed out some years ago, our field of applied linguistics is much better at analysis than at synthesis.

Listen and give the right answer tasks. Here, I include work being done in the area of standardized listening tests–the TOEFL-type listening comprehension format, where, although the length and complexity of the stimulus vary, the main task is to listen and then answer questions about the text. The research and development work in this category have recently been focused on item response theory (e.g., Woods & Baker, 1985) and the problems of developing equivalent forms and of adapting items to individual candidates' proficiency levels. As reported by Stansfield (1986), work has also been progressing in the TOEFL program on investigating the effects of manipulating the stimulus material by including such features as pauses, expressions of hesitation and uncertainty, false starts, and appropriate background noises (e.g., talk and clanking dishes in a restaurant situation).

Reduced redundancy tasks. In these tests, the primary focus is to give candidates an opportunity to show that they can process language in spite of reduced redundancy, effected usually by mutilating the text in some way or by adding "noise" to the transmission. Candidates may demonstrate proficiency either by giving the right answer in spite of the noise or mutilation or by filling in the missing or masked portions of the text. The most well-known examples of reduced-redundancy tests include; a) the *noise test* (Gradman & Spolsky, 1975), produced by adding white noise to the transmission of stimulus sentences, which candidates have to reproduce in writing; and b) partial dictation and listening cloze tests (Henning, Gary & Gary, 1983), where candidates are given a mutilated written text and have to fill in the gaps by listening to an unmutilated spoken version.

Repetition/imitation tasks. Here I would include: a) work on dictation tests (Fouly & Cziko, 1985), particularly the recent work of

Cziko and his colleagues in studying the scalability of dictation-test results, based on chunks of increasing length; b) the Integrative Grammar Test (Bowen, 1976), an interesting example of the use of natural reduced redundancy (reduction, assimilation and contraction) to test grammar; and c) elicited repetition (Perkins, Brutton & Angelis, 1986), where students are asked to repeat sentences as a demonstration of their control over specific syntactic structures. Mydlarski and Paramskas (1985) have developed a computerized template program to allow teachers easily to develop dictation exercises and tests in which students respond via the keyboard.

Interaction tasks. In tests involving interaction tasks, candidates demonstrate comprehension by providing an appropriate response to a conversational stimulus. Formats of interaction tests range from oral interviews, which frequently make provision for a comprehension score, to multiple-choice formats where candidates are asked to choose an appropriate response. One interesting interaction test, reported by Friel (1984), is called the *Discourse Cooperation Test*, which is based on Grice's maxims of conversational cooperation. In the test, a situation is established, such as an automobile accident, and the examiner initiates the interaction by flouting one or another of the Gricean maxims. The candidate demonstrates comprehension of the flouted maxim by choosing an appropriate response from five options, four of which are each related to one of the maxims, plus one neutral response.

Computerized interactive audio/video programs, such as the *Montevidisco* project at Brigham Young University (Gale, 1983; Schneider & Bennion, 1983) also offer an interactive listening-comprehension testing capability. For example, the student is confronted, by means of videodisc, by a native of a Mexican village and asked whether he or she would like a tour guide. The student demonstrates comprehension by choosing an appropriate response, which may include a request for repetition (which, owing to the magic of the videodisc technology, is readily supplied). Interactive audio/video applications will be discussed in more detail in the *Technology* section.

Media-transfer tasks. Here, I would place a variety of listening comprehension test activities that involve the candidates' abilities to transfer between two or more media–aural, written, pictorial, diagrammatic–to demonstrate comprehension. For example, the Michigan English Language Battery Listening Comprehension subtest presents students with a map of a picnic area while providing information aurally for the interpretation of the map. Comprehension is shown by the students' ability to answer questions about the map. The British Royal Society of Arts examination requires students to complete a short schedule by listening to an announcement; The British Joint Matriculation Board examination has a section where students hear an explanation of the technical labels on refinery gas lines and

then have to fill in blank labels based on further aural information. Also in Britain, the new Test in English for Educational Purposes (Emmett, 1985) involves candidates in a preliminary reading task, followed by a related lecture on which comprehension questions are asked. Candidates then write down some dictated notes and complete a writing task based on what they have read and heard. As a final example of this integrated transfer type of test, Madden and Samuda (1985) report on a task-based test in which candidates listen to a conversation between two people planning their day's activities. The students have a map on which they are to mark the route of the planned excursion, and a small table to fill in with information about times and places. This is followed by reading, writing, and speaking tasks, all integrated around not only the topic but previous input.

This brief review suggests that current work in listening comprehension test development is moving in the following directions:

1. Tests are more efficient and adaptable to individual abilities.
2. Tests are both integrative (in the technical sense of testing more than one skill or ability at once) and integrated (in the sense of listening comprehension tasks feeding and being fed by other skills).
3. Tests challenge learners to employ varying strategies with varying conditions of transmission noise, text mutilation, or differing media.
4. Tests are contextualized.
5. Tests require students to apply understanding gained through listening to a variety of tasks involving other components of language proficiency.

In light of this understanding of current work, I would like to turn now to a consideration of the *ACTFL Proficiency Guidelines* and suggest directions for test development in the ACTFL context, particularly those that came out of a symposium and working group on testing the receptive skills reported in the September, 1984, issue of *Foreign Language Annals*.

Testing the Receptive Skills

Canale (1984) has suggested two overriding principles in considering the testing of reading and listening proficiency: a) attempt to elicit the best performance from test takers by presenting tasks that are fair, important, and interesting in themselves; and b) attempt to provide test tasks that can reveal to test takers and educators clear,

rich, relevant, and generalizable information. More detail was provided in a synthesis produced by Canale, Child, Jones, Liskin-Gasparro, and Lowe (1984) in which the following recommendations were made:

> 1. A focus should be placed on criterion-referenced proficiency testing that would provide both profile scores and fine-grained diagnostic information for a variety of purposes.
> 2. Test organization should be based on thematic unity (as opposed to coverage simply of structural points or skill areas) and the four ILR oral interview stages (warm-up, level-check, probe, and wind-down).
> 3. Item specifications should reflect the interrelationship among performance level, text type, reading/listening function, and reading/listening strategy (See the Appendix for a table outlining this interrelationship).
> 4. An adaptive administration procedure should be devised to provide an efficient, accurate, and affectively rewarding test administration.
> 5. Scoring criteria should be accuracy of comprehension, independence in dealing with text, and speed of completion.
> 6. Item difficulty should be calibrated by means of reference to proficiency levels, pilot screening, and on-going studies. Construct-, predictive-, concurrent-, and content-validity studies should be carried out.
> 7. Learners, teachers, and administrators should each receive useful feedback from the tests, such as raters' commentary and verbal summaries of group and individual performance.

These recommendations are reviewed by Dandonoli (1987), who then goes on to describe current reading and listening test development within ACTFL. She provides an interesting and thoughtful discussion of the problems likely to be encountered in carrying out the recommendations of the 1984 working group, and her views will be referred to frequently in the remainder of this paper.

The current *ACTFL Proficiency Guidelines* (1987), "identify stages of proficiency...to allow assessment of what an individual can do and cannot do with language, regardless of where, when, or how the language has been learned or acquired" (p. 15). Five levels are identified: novice, intermediate, advanced, superior and distinguished.

The first two stages have low, mid and high substages, while the third stage has an advanced-plus category. The generic description of the listening comprehension guidelines assumes "that all listening tasks take place in an authentic environment at normal rate of speech using standard or near-standard forms" (p. 18). The novice level is characterized by topics referring to basic personal information or the immediate physical surroundings; intermediate, by topics on personal and social conventions; advanced, by topics beyond the immediacy of the situation dealing with factual information; the superior level, by abstract topics in professional settings; and the distinguished level, by comprehension of material on any topic, for a variety of needs tailored to different audiences.

Though it is beyond the scope of this paper to critique the current form of the *Guidelines*, I would mention some surprise that there is not closer adherence to the text and task specification outline provided by Canale and his colleagues (1984, reproduced here in the Appendix); I find it hard to discern the supposed developmental relationship among the five stages of the *Guidelines*. Apart from that criticism, which may be one merely of style and rhetoric, we must continue to remind ourselves that if the *ACTFL Proficiency Guidelines* are to be truly said to identify stages of proficiency, empirical studies must be conducted to show, for example, that learners who can "understand main ideas and some facts from interactive exchanges" can also "recognize learned material and isolated words and phrases when strongly supported by context" (p. 18), but cannot do these things "beyond the immediacy of the situation" (p. 19), and so on up the scale. In other words, it is crucial that we obtain satisfactory information on the validity of the underlying construct of the *Guidelines* (cf. Savignon, 1985) before and in concert with further work on a national metric for foreign language proficiency.

Before turning to a detailed discussion of specific areas that seem to me to be both particularly interesting and nettlesome, I would like to draw attention to the very lively debate on the *Proficiency Guidelines* and their impact that has been reported in various journals in the last couple of years (Bachman & Savignon, 1986; Kramsch, 1986 and 1987; Lantolf & Frawley, 1985; Lowe, 1986; Savignon, 1985; Schulz, 1986). The discussion has been undertaken to refine the concept of proficiency from a communicative, interactionist point of view. The current *Guidelines* already reflect some of the points made in the debate, and no doubt other points have emerged from the Symposium on the Evaluation of Foreign Language Proficiency (Bloomington, Indiana, 1987). Interestingly, the result of this refining process has been not a narrowing of the notion of proficiency by winnowing out unnecessary components, but rather a widening of it to accommodate our growing

understanding of how complex language in use is and how varied are the purposes for which we employ it.

Issues for Discussion

I would now like to take up, in turn, four issues that seem to me to be of particular importance in our consideration of listening comprehension testing in the context of current research, test development and the *ACTFL Proficiency Guidelines*: criterion-referenced testing, context in listening comprehension, language for specific purposes, and uses of technology.

Criterion-Referenced Testing

We have seen (Canale et al., 1984; Dandonoli, 1987) that research concerning the testing of receptive skills in the ACTFL context involves the concept of criterion referenced testing. I am frankly uneasy about the meaning of *criterion reference* in a language testing enterprise. The perceived advantage of criterion-reference is succinctly stated by Delamere (1985): "Criterion-referenced tests make it possible for us to draw different kinds of inferences about a student's ability to use the language than norm-referenced tests do...For example, a criterion-referenced test might enable us to say that an individual learner spelled eight out of ten words correctly from a basic sixth grade vocabulary list and that his performance is therefore satisfactory" (p. 45). Such a view highlights what I see as a flaw in the criterion-referenced position with regard to language testing, namely, that it is beyond our present understanding of how language is structured and used to be able to say with any confidence that because a learner spelled eight out of ten words correctly his performance is, therefore, satisfactory. Or alternatively, that because a learner spelled only seven out of ten words correctly, his performance is unsatisfactory. As Denham (1975) points out, the establishment of a criterion level is often quite arbitrary; it may be that Cziko's term *edumetric* (Cziko, 1983) offers a solution to this problem: He would link linguistic criteria to educational goals. However, it seems to me that the real difficulty is one reflecting the nature of language use itself: There are so many ways, structural and strategic, to "get something done" with language that it is currently beyond our ability to establish a domain of tasks which would add up to a criterion. Only a very narrowly focused language for specific purposes context (cf. Widdowson, 1983) might allow for a reasonable criterion referencing. Cartier (1968) illustrates how this might be possible where "we know exactly where the student will go, what he will be studying

there, and what kind of work he will be doing afterward" (p. 29). However, the *ACTFL Guidelines* are directed at much more general purposes and, at our present state of linguistic understanding, an attempt at criterion-referenced language testing will necessarily be reductionist. As Lantolf and Frawley (1985) argue, "such tests construe A world; they do not reflect THE world" (p. 343). As Tollefson (1986) and Auerbach (1986) warn, a number of limiting biases are likely in materials and tests developed with explicit competencies in mind, selected, as they are, without an accepted methodology for discovering essential competencies. Discussion needs to focus on the meaning of criterion referencing in language testing and on what kind of world our criteria might impose on learners.

Context and Listening Tests

Canale et al. (1984) have recommended that tests be organized around a unifying theme as a means of providing a more satisfying and motivating experience for candidates. Dandonoli (1987) supports this and points out that thematic unity also has the advantage of providing an authentic and meaningful context. Discussion needs to focus on what we mean exactly by *thematic unity* and by *context* and also on how context influences language performance. Brown (1986), for example, distinguishes between *external context*, which constrains expectations for what is appropriate, and *discourse-internal context*, which establishes shared information and conventions. She claims that discourse-internal context is harder for less able students to recognize and use. The use of internal context involves an "ability to extrapolate from previous moves in the task and so to predict the expected range of the next move" (p. 300). Brown reports research on first-language listening comprehension in which subjects shared information on maps, and suggests that good and poor listeners differed not so much in their abilities to recognize inadequate messages, but more in "the ability to give appropriate feedback to the speaker when the message is inadequate" (p.299). A refined notion is needed of what in the context is relevant for communication.

Faerch and Kasper (1986) list some fairly uncontroversial hypotheses in current models of first language listening comprehension, including the following:

> 1) Comprehension processes rely on three types of information: linguistic input, recipient's linguistic and sociocultural knowledge, and contextual information.
> 2) The process of matching input with the recipient's knowledge can start either with the input or with the

knowledge. In both cases, non-textual information is used to support comprehension.
3) Different comprehension tasks focus on one or the other of the two processing directions.
4) It is the rule rather than the exception that the match between input and knowledge is less than perfect, resulting in gaps that indicate comprehension is typically partial rather than total.

Faerch and Kasper suggest that second language comprehension differs from first language comprehension only in that second language learners have more restricted knowledge of the second language, influenced by transfer from the first language, and have unrealistic expectations that comprehension must be total.

Overall, research such as that by Brown and Faerch and Kasper suggests that the influence of context on second language comprehension, although acknowledged to be important, is little understood as yet. Indeed, the work of such researchers as Tarone (1983) suggests that different tasks and/or contexts result in different second language performances by learners. Sotillo (1987), in a case study of a *fossilized* learner who had been in the United States for 20 years and who is currently working in an urban hospital, concluded that a domain-specific English language instructional program, based on an analysis of the social and linguistic needs for interacting with patients in the hospital environment, yielded greater interlanguage destabilization and growth toward a target *American* norm than did a traditional ESL program, based on analysis of grammatical and lexical features of spoken and written English. This work suggests that language may be learned in domain-specific contexts. We need to consider what this might suggest for how students ought to be tested. Much more information is needed in this area of proficiency.

Language for Specific Purposes

Although *Proficiency* is defined as the ability to use language for some real-life purpose (Dandonoli, 1987, p. 78), one finds surprisingly little in the *Guidelines* about the specific purposes of language use around which the testing enterprise can be built. Discussion should focus, I think, on the issue of purpose, and how related issues of criterion referencing, thematic unity, and context might be dealt with. The table illustrating text and task specifications for reading and listening, provided by Canale et al. (1984), for example, related to work by Child (1987), is devoid of any reference to specific domains of language use. Somehow, it seems to me, provision must be made for incorporating the

concept of specific purposes into the *general purpose* ACTFL context. The work of Widdowson (1983) may be a starting place for discussion of this area.

Widdowson (1983) suggests that a difference between specific- purpose and general-purpose language teaching lies in the distinction between *purpose* as a descriptive term in specific-purpose operations and as a theoretical term in general-purpose ones. What he means is that in general purpose teaching, where the eventual use of the language being learned is not clearly discernable (the case being addressed by the *ACTFL Guidelines)*, "purposes are specified by objectives, pedagogical constructs which seek to provide for the achievement of communicative aims when occasion arises after completion of the course" (p.12). He would have us focus, in our specification of purposes of language use, on procedures for *making sense*–procedures being defined as negotiating activities which interpret the directions provided and enable us to alter our expectations in the light of new evidence as the discourse proceeds" (pp.40-41). How do we arrive at an understanding of these procedures so as to include them in our teaching and testing? Widdowson suggests a problem-oriented approach to activating procedures, providing learners with the same type of procedural work as would be required in their various activities outside the language classroom (p.91). Thus it seems to me that procedural activity might legitimately be a focus of listening comprehension test development, testing subjects' abilities to *make sense* of material related to specified domains of language use.

The problem that needs discussion, I believe, is the specification of domains of language use in the ACTFL context. What purposes (in the Widdowson *procedural* sense) do we have in mind when we speak of testing proficiency? *Proficiency* for what? It will not be enough to speak of thematic unity, of criterion referencing or of text/task specification, without relating these notions to procedural activity associated with specific domains of language use.

Technology

There are two issues related to the use of technology in listening comprehension testing that might fruitfully be discussed: that of computer adaptive testing (CAT) employing item response theory, and that of interactive audio/video (IAV) techniques.

Computer adaptive testing. Dandonoli (1987) has suggested that computer adaptive testing procedures merit consideration in test development within the context of the *ACTFL Guidelines*, owing mainly, in her view, to the promise of greater efficiency. She cautions that there are some theoretical difficulties to be overcome, most notably

that of the unidimensionality requirement of item response theory as well as that of a potential narrowing of the content range of a test owing to a reduction in the number of items allowed by the application of item response theory. Larson and Madsen (1985) argue that computer adaptive tests provide a number of advantages to the teacher, tester, and student; among them are: its compatibility with computer assisted instruction, immediate feedback of results, ease of revision of test content, individualizing test items, and flexibility in test length. They are particularly interested in the possibilities in item response theory, first, of presenting items within a range of difficulty close to the ability level of the examinee as a method of increasing test effectiveness, and second, of placing examinees on an absolute scale of proficiency, independent of the persons being measured and of the measure being used.

Canale (1986), in writing of the "promise and threat" of computerized adaptive testing, attempts a balanced view of the advantages and problems, and in particular cautions that we may be in danger of trivializing our theories and of compromising their relevance to educational settings (p.33). This echoes Tung (1986), who also cautions that the requirement of unidimensionality in item response theory "raises serious doubts whether complex language behavior characteristic of real life communication can be assessed with CAT" (p.25). Canale also raises the issue of method effect in the use of CAT and calls for research into the potential problems associated with the mechanization of comprehension testing. Useful studies might be carried out comparing tests constructed outside the limits of item response theory with those adhering to its requirements. What desirable features of content are excluded or limited by item response theory, and how might these be incorporated in adaptive tests constructed by other means?

Interactive audio/video. The potential offered through the interfacing of the computer, videodisc, and audio recorder, referred to as interactive audio/video, is almost overpowering. The *Montevidisco* project (Gale, 1983; Schneider & Bennion, 1983) has already been mentioned. It takes students on a simulated visit to a northern Mexican town and in the process exposes them to real-life situations, with natives speaking in Spanish. There are slightly different versions for male and female students, owing to the gender intensive nature of Spanish. The overall effect is rather like that of a computer adventure game in which the player controls movement through a series of scenes and situations and must produce appropriate responses to them. In *Montevidisco* there are 28 major sequences, each with several scenes, each scene having at least four options for the student to select. In all there are over 1100 possible paths, and the program can be run a number of times, with the student encountering novel situations each

time. Rowe (1985) sees IAV in the style of *Montevidisco* as the wave of the future, and calls for a NASA-like cooperative development effort "to deliver the goods" in an IAV mode. Other programs employing IAV include *Klavier im Haus* (Schrupp, Bush, & Mueller, 1983), a short program in German, following the problems of a couple moving into a new apartment; *Videodeutch* (Dahl & Luckau,1985), another program in German which emphasizes nonverbal cultural aspects of language learning; and a pilot videodisc project in French (Eastmond & Mosenthal, 1985), which builds on the *Montevidisco* project by allowing student access to a main menu at any time to ask for help, change sound tracks from the original dialogue to one with a clear narrative, call up subtitles in either their first language or French, or call up a complete written text with translation.

What IAV offers the tester of listening is a capacity for the student to demonstrate comprehension by responding appropriately to natural, contextualized stimuli in the program by entering a multiple choice response on the keyboard, or by recording it orally on the cassette. The computer can easily make a diagnostic record of the responses, including requests for repetition and the speed with which responses were made.

One should also mention the somewhat more one-way listening possibilities offered by such technical innovations as connecting audio-cassettes to computers (e.g., PLATO IV at the University of Illinois) and digitally recorded voice data on floppy disks. Both types of audio material can be part of the computer program and can be randomly accessed. Listening cloze, notetaking, text reconstruction and deductive problem solving tests can all be devised using the interaction of the computer and an audio device.

What needs discussion, of course, are the potential disadvantages of pursuing an IAV approach to language teaching/testing. They include those already mentioned in connection with computer adaptive testing, since presumably one might want to employ CAT techniques in the construction of the IAV program. The question of method effect is particularly relevant here, for it is little understood how students might be affected when confronted by the specter of videodisc player, audiocassette recorder, computer, keyboard, monitor, TV screen, microphone, and headphones. In what sense can they be said to be engaged in natural communication? But there are also problems of the enormous cost and relative unavailability of IAV systems: The commercial viability of the videodisc has yet to be proven (indeed, one original manufacturer of the technology has already discontinued production). One project, in the face of this sort of problem, has moved to the use of the videocassette in interactive programs: *Autotutor* (Little & Davis, 1986) has the advantage of relatively low cost and the availability of huge amounts of video material adaptable to language

teaching and testing. What it doesn't allow, of course, is immediate access to any part of the program, which is an essential component of interaction. Research reported by Schrupp, Bush, and Mueller (1983) suggests that without the interactive element, video loses some of its power as a teaching medium. It may be that we await further technological advances which will provide for the interactive capability of the videodisc without its expense and relative unavailability.

Conclusion

In this concluding section, I want to summarize suggestions for development and validation studies in listening comprehension testing. Such research and development would be toward tests that reflect our understanding of the listening skill as complex and active. We seem to be moving in the direction of, first, tests that are more efficient in the sense of being more adaptable to individual abilities; second, tests that are both integrative, in the technical sense of testing more than one skill at a time, and integrated in the sense of the listening skill being fed by and feeding other skills; third, tests that challenge learners to employ strategies to deal with varying conditions of noise, text mutilation, or media; fourth, tests that are contextualized; and fifth, tests that require learners to apply understanding gained through listening to a variety of tasks. Proficiency in listening comprehension would be related to empirical stages of second language development by reference to scales reflecting such development. Finally, test formats would be derived from technology associated with computer adaptive testing and interactive audio/video techniques. Validation studies will focus on relating test results with developmental stages, and on analyzing test content in relation to our understanding of the construct of language proficiency.

In considering these tendencies in applied linguistics and language testing within the context of the *ACTFL Proficiency Guidelines* and the movement toward a national language proficiency metric, I outlined four issues that seem to me to be both extremely interesting and thorny, and which might profitably be discussed: criterion-referenced testing, context and listening tests, language for specific purposes, and uses of technology in testing.

Concerning the first issue, criterion referencing, we need, I think, to focus on problems arising from the nature of language itself–form, and function relationships are just not as clear as we might once have hoped. It may be that we would want to consider the advantages of interpreting language proficiency in terms of domains of behavior rather than in terms of criterion levels (Denham, 1975), but we need to consider carefully the effect of the resulting reductionism. To repeat the

point made by Lantolf and Frawley (1985): criterion referenced tests "construe A world; they do not reflect THE world" (p. 343). Clark and Clifford (1987) have called for validation studies relating the test results to highly detailed, task-specific external criteria. It seems to me that it is precisely those highly detailed task-specific criteria that ought to form the basis of the tests themselves. The results would be a sampling of content domains and a more truly criterion-referenced set of measures.

The question of context raises a number of interrelated issues including thematic unity, the effect of background knowledge on test results (the hot house special problem), the effect of topic on variation in learner language output, and the effect of internal context on comprehension. We must come to grips with exactly what it means to speak of contextualized language tests. In terms of research, the question of test-retest, and intra/interrater reliability studies seems relevant here. Claire Kramsch (personal communication) has questioned whether test raters might not be socialized into being good raters. I would pose the question differently: What is the construct by which raters arrive at a rating? Do raters arrive at similar ratings for different reasons? We know that it is the nature of language itself that there are many ways to get the same job done. The problem we need to explore is not only the possiblity that learners with similar competence might be given different scores on our tests, but also that learners who make similar scores might do so as a result of different competence. Rater-related constructs as well as learner-related ones are of crucial importance to our understanding of variation in language production in context.

What specific domains of language use will the *ACTFL Guidelines* and tests be directed toward? We need to go beyond the relatively simple notion of thematic unity, and focus attention on purpose, or in Widdowson's (1983) terms, procedural activity. Widdowson proposes a focus on procedures partly as a way out of a specificity dilemma–the logical absurdity of idiosyncratic teaching and testing materials that many see resulting from the Language for Specific Purpose approach. We must consider ways of testing learners' abilities to make sense of linguistic input relative to specific domains of use and answer the question, "Proficiency for what?" It may be that we can learn from the experience of the Council of Europe's Modern Languages Project, which has established *waystage* and *threshold* level specifications based upon "communicative needs" for syllabus and test development. Trim (1984) captures an essential feature of the Council of Europe approach: "A learning biography consists not of a straight-line progress from elementary to intermediate to advanced, but an accumulation of life-related learning experiences" (p.20). Obviously, our thoughts on criterion- versus domain-referencing may be relevant here, but the

question of purpose in listening tests must remain a key one. This European perspective, encapsulated in the Trim quotation, is fundamentally different from that implied in the *ACTFL Guidelines*. It is that language is learned socially, purposefully, and cumulatively in life experience. It follows, the Council of Europe team suggests, that language should be taught and tested experientially, in life-related units. Liskin-Gasparro (1987) has referred to the difficulty of incorporating culture into the *Proficiency Guidelines*, and it has been the experience of examiners in the ILR oral interview (Pardee Lowe, personal communication) that culture becomes an issue only somewhere around level three. At the same time, as Clark and Clifford (1987) indicate, "the performance of students within most regular academic settings (including even college language majors; see Carroll, 1967) would not be expected to go beyond level 2 or 2+" (p. 4). Can it be that culture need not or cannot be taken into account in our proficiency tests in academic programs?

Gass (1987) has suggested that we need to take into account the way language is learned as we develop tests and guidelines. In this context, we need to remember that language is learned in culture and in life experience. Selinker and Douglas (1985) have used the term "slices of life" to refer to the internal contexts within which, they hypothesize, interlanguages are constructed. It may be that the insistence upon viewing language learning linearly–from elementary to intermediate to advanced–distracts us from an essential fact: language learning is not linear; language is learned in units of experience. It seems to me that we cannot produce tests that represent a linear view of learning and then try to see how they are correlated to external criteria within life experiences. The purposeful, life-related elements must be there from the start.

Regarding technology in listening comprehension testing, I noted two issues: computer adaptive testing and the use of interactive audio/video techniques. We need to reflect upon CAT in terms suggested by Canale (1986), who calls for studies of adaptive tests constructed outside the possibly restrictive psychometric limits of item response theory in order to determine what essential features of communicative language proficiency might be lost by adhering to those restrictions. The question of IAV is primarily one of the interface between learner and machine, and its effect on language performance–both what technology does to language and what it does to learners (e.g., Carol Chapelle, at Iowa State University, is studying the relationship between various learner characteristics and computer assisted language learning).

One might suggest (only somewhat tongue-in-cheek) that we are searching for a Grand Unification Theory in language testing that will relate constructs associated with language competence, domains of use,

stages of learning, psychometric requirements, and technological potential. The continuing search will be worth participating in.

References

American Council on the Teaching of Foreign Languages. (1987). In H. Byrnes & M. Canale (Eds.), *Defining and developing proficiency: Guidelines, implementations and concepts*. The ACTFL Foreign Language Education Series (pp. 15-24). Lincolnwood, IL: National Textbook.

Auerbach, E. (1986). Competency-based ESL: One step forward or two steps back? *TESOL, Quarterly, 20*, 411-429.

Bachman, L. & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal, 7*, 380-390.

Bowen, J. D. (1976). Current research on an integrative test of English grammar. *Regional English Language Center Journal, 7*, 30-37.

Brown, G. (1986). Investigating listening comprehension in context. Applied *Linguistics, 7*, 284-302.

Canale, M. (1984). Considerations in the testing of reading and listening proficiency. *Foreign Language Annals, 17*, 349-357.

Canale, M. (1986). The promise and threat of computerized adaptive assessment of reading comprehension. In C. Stansfield (Ed.), *Technology and language testing* (pp.29-46). Washington, DC: TESOL.

Canale, M., Child, J., Jones, R., Liskin-Gasparro, J. & Lowe, P. (1984). The testing of reading and listening proficiency: A synthesis. *Foreign Language Annals, 17*, 369-392.

Carroll, J. B. (1967). *The foreign language attainments of language majors in the senior year: A survey conducted in U.S.colleges and universities*. Cambridge, MA: Harvard University Graduate School of Education.

Cartier, F. A. (1968). Criterion-referenced testing of language skills. *TESOL Quarterly, 2*, 27-32.

Chaudron, C. & Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics, 7*, 113-127,

Child, J. R. (1987). Language proficiency levels and the typology of texts. In H. Byrnes and M. Canale (Eds.), *Defining and developing proficiency: Guidelines, implementations and concepts* (pp. 97-104). Lincolnwood, IL: National Textbook.

Clark, J. & Clifford, R. (1987). The FSI./ILR/ACTFL proficiency scales and testing techniques: Development, current status and needed research. Paper presented at the symposium on the evaluation of foreign language proficiency, Indiana University, Bloomington.

Dziko, G. (1983). Psychometric and edumetric approaches to language testing. In J. W. Oller (Ed.). *Issues in language testing research* (pp. 289-307). Rowley, MA: Newbury House.

Dahl, R.C. & Luckau, P. (1985). Videodeutch: a computer assisted approach to verbal and non-verbal cultural literacy. *CALICO Journal: 2* (4), 13-19.

Dandonoli, P. (1987). ACTFL's current research in proficiency testing. In *Defining and developing proficiency* (pp. 75-96). See Child 1987.

Delamere, T. (1985). Notional-functional syllabi and criterion-referenced tests: The missing link. *System. 13*, 43-47.

Denham, C. H. (1975). Criterion-referenced, domain-referenced and norm-referenced measurement: A paralax view. *Educational Technology, 15*, 9-13.

Douglas, D. & Pettinari C. (1982). Grounded ethnography: A method adaptable to the production of communicative ESP materials. Paper presented at TESOL Convention, Toronto.

Dunkel, P. (1986). Developing listening fluency in L2: Theoretical considerations. *Modern Language Journal, 70*, 99-106.

Eastmond, J. N. & Mosenthal, R. (1985). The World Center for Computing's pilot videodisc project for French language instruction. *CALICO Journal, 2* (4), 8-12.

Emmett, A. (1985). Test in English for Educational Purposes (TEEP) - a new test from the Associated Examining Board (AEB). *Language Testing Update, 1*, 3.

Faerch, C. & Kasper, G. (1986). The role of comprehension in second-language learning. *Applied Linguistics, 7*, 257-274.

Fouly, K. A. & Cziko, G. (1985). Determining the reliability, validity and scalability of the graduated dictation test. *Language Learning, 35*, 555-566.

Friel, M. (1984). The discourse co-operation test. *System, 12*, 251-262.

Gale, L. E. (1983). Montevidisco: An anecdotal history of an interactive videodisc. *CALICO Journal, 1*, 42-46..

Gary, J. & Gary, N. (1981). Caution: Talking may be dangerous to your linguistic health. *IRAL, 19*, 1-14..

Gass, S. (1987). L2 vocabulary acquisition. Paper presented at the Symposium on the Evaluation of Foreign Language Proficiency, Indiana University, Bloomington.

Glisan, E. W. (1985). The effect of word order on listeningcomprehension and pattern retention: An experiment in Spanish as a foreign language. *Language Learning, 35*, 443-469.

Gradman, H. & Spolsky, B. (1975). Reduced redundancy testing: A progress report. In R. Jones and B. Spolsky (Eds.). *Testing language proficiency* (pp. 59-70). Arlington, VA: Center for Applied Linguistics.

Henning, G., Gary, N., & Gary, J. (1983). Listening recall–a listening comprehension test for low proficiency learners. *System, 11*, 287-293.

Kramsch, C. (1986). From language proficiency to interactional competence. *Modern Language Journal, 70*, 366-371.

Kramsch, C. (1987). The proficiency movement: A second language acquisition perspective. *Studies in Second Language Acquisition, 9*, 355-362.

Lantolf, J. & Frawley, W. (1985). OPT: A critical analysis. *Modern Language Journal, 69*, 337-345.

Larson, J. & Madsen, H. (1985). Computerized adaptive language testing: Moving beyond computer-assisted testing. *CALICO Journal, 2* (3), 32-36.

Liskin-Gasparro, J. (1987). The ACTFL proficiency guidelines: An update. Paper presented at the symposium on the Evaluation of Foreign Language Proficiency, Indiana University, Bloomington.

Little, D. & Davis, E. (1986). Interactive video for language learning: The autotutor project. *System, 14*, 29-34.

Lowe, P. (1986). Proficiency: Panacea, framework, process? A reply to Kramsch, Schulz and particularly to Bachman and Savignon. *Modern Language Journal, 70*, 391-397.

Madden, C. G., & Samuda, V. (1985). Task-based test design. *Language Testing Update, 1*, 11-15.

Mydlarski, D. & Paramskas, D. (1985). Template system for second language aural comprehension. *CALICO Journal, 3* (2), 8-12.

Perkins, K., Brutton, S, & Angelis, P. (1986). Derivational complexity and item difficulty in a sentence repetition task. *Language Learning, 36*, 125-141.

Powers, D. E. (1986). Academic demands related to listening skills. *Language Testing, 3*, 1-38.

Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*, 219-240.

Rowe, A. A. (1985). Interactive language simulation systems: Technology for a national language base. *CALICO Journal, 2*, 44-47.

Savignon, S. (1985). Evaluation of communicative competence: The ACTFL provisional proficiency guidelines. *Modern Language Journal, 69*, 129-134.

Schneider, E.W. & Bennion, J. (1983). Veni, vidi, vici via videodisc: A simulator for instructional conversations. *System, 11*, 41-46.

Schrupp, D. M., Bush, M. & Mueller, G. (1983). Klavier im haus–an interactive experiment in foreign language instruction. *CALICO Journal, 1* (2), 17-21.

Schulz, R. (1986). From achievement to proficiency through classroom instruction: Some caveats. *Modern Language Journal, 70*, 373-379.

Selinker, L. & Douglas, D. (1985). Wrestling with "context" in interlanguage studies. *Applied Linguistics, 6*, 190-204.

Sotillo, S. (1987). The impact of domain-specific instruction on language use: The case of a long-term resident. Paper presented at American Association of Applied Linguistics Conference, San Francisco.

Stansfield, C. W. (1986). Introduction. In C. W. Stansfield (Ed.), *Toward communicative competence testing: Proceedings of the Second TOEFL Invitational Conference* (pp. ix-xv). TOEFL Research Report 21. Princeton: Educational Testing Service.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142-163..

Tollefson, J. W. (1986). Functional competencies in the U.S. refugee program: Theoretical and practical problems. *TESOL Quarterly, 20*, 649-664.

Trim, J. L. M. (1984). Extract from *Developing a unit/credit scheme of adult language learning*. Reprinted in J. A. Van Ek & J. L .N. Trim (Eds.), *Across the threshold* (pp. 9-26). Oxford: Pergamon.

Tung, P. (1986). Computerized adaptive testing: Implications for language test developers. In C. W. Stansfield (Ed.), *Technology and language testing* (pp. 11-28). Washington, DC: TESOL.

Tyler, A. (1987). Sources of the perception of incoherence in non-native spoken discourse. Paper presented at American Association of Applied Linguistics Conference, San Francisco.

Widdowson, H. (1983). Learning purpose and language use. Oxford: Oxford University Press.

Woods, A. & Baker, R. (1985). Item Response Theory. *Language Testing*, *2*, 119-140.

APPENDIX .

Text and Task Specifications for Reading (R) and Listening (L)

Level	Text Type	R/L Function	R/L Strategy
0/0 +	Enumerative (numbers, names, street signs, isolated words, phrases)	Discern discrete elements	Recognize memorized elements
1	Orientational (simplest connected text, such as concrete descriptions)	Identify main idea	Skim, scan
2	Instructive (simple authentic text in familiar contexts and in predictable sequence)	Understand facts, literal information	Decode, classify
3	Evaluative (authentic text on unfamiliar topics)	Grasp ideas, implications	Infer, guess hypothesize, interpret
4	Projective (all styles and forms of text for professional needs or for general public)	Deal with unpredictability, cultural references	Analyze, verify, extend hypotheses
5	Special purposes (extremely difficult and abstract text)	Equivalent to educated native reader/listener	All those used by educated native reader/ listener

From Canale et al. (1984), p. 391.

Author

DAN DOUGLAS, Assistant Professor, Department of English, 203 Ross Hall, Iowa State University, Ames, IA 50011. Specializations: applied linguistics, second language acquisition, language testing.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editors encourage the submission of manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to Editors, *Applied Language Learning*, Faculty Development Division, Defense Language Institute Foreign Language Center, Presidio of Monterey, CA 93944-5006.

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on $8\frac{1}{2}$ x 11 inch paper, double-spaced, with ample margins. Where feasible, manuscripts are preferred on disc (DOS or XEROX Viewpoint). Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and if necessary notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

Review Process

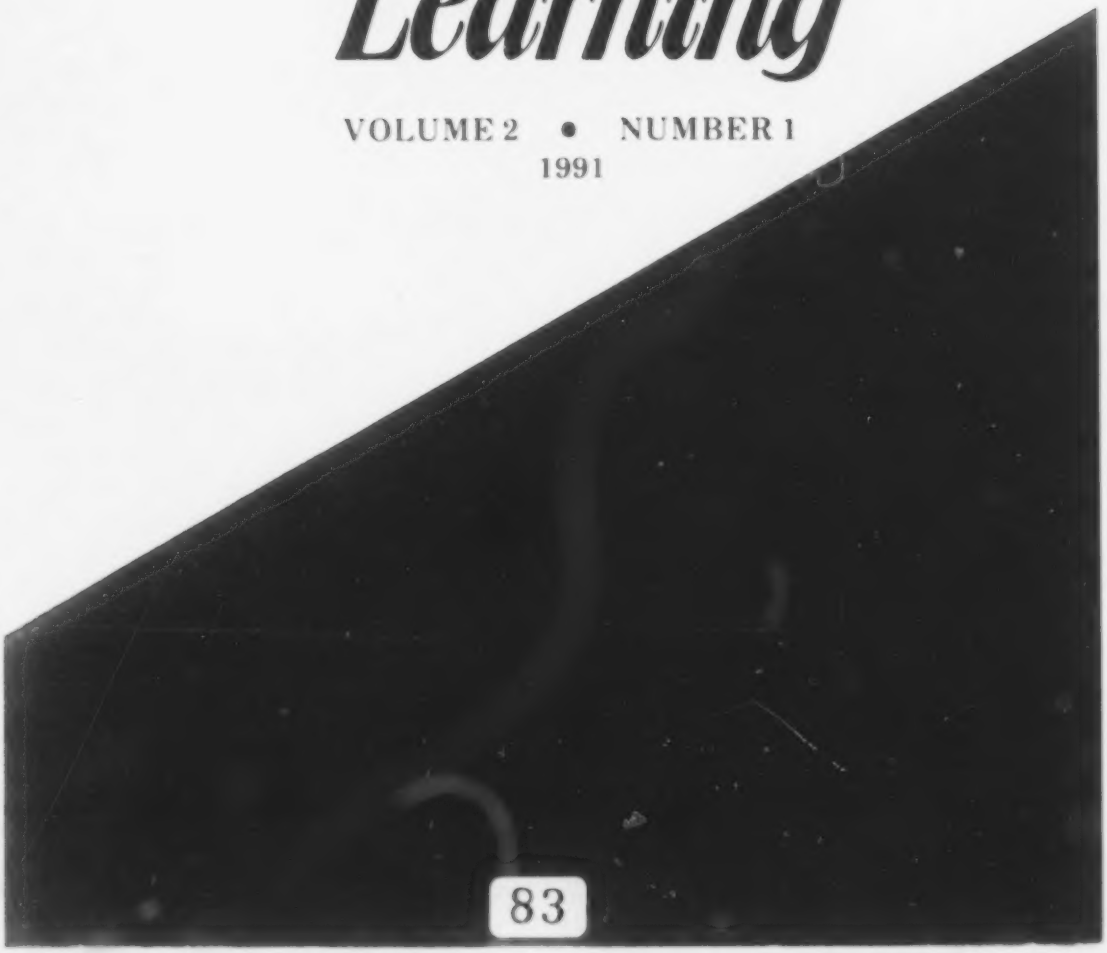
Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

Applied Language Learning

VOLUME 2 • NUMBER 1
1991



83

Applied Language Learning

Volume 2 • Number 1

1991

Defense Language Institute
Foreign Language Center

APPLIED LANGUAGE LEARNING

PB 65 - 91 - 1

1991 • Volume 2 • Number 1

Applied Language Learning (ISSN 1041-679X) is published semi-annually by the Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. *ALL* presents critical integrative articles and research reports on adult language learning for functional purposes.

COL Donald C. Fisher, Jr., Commandant
Defense Language Institute, Foreign Language Center
Department of the Army

This publication presents professional information: the views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material to meet space constraints.

This publication approved for public release: distribution is unlimited.

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the Institute. Cover design and graphics by Barney Inada, Visual Production, DLI; wordprocessing, Virginia Wilson, Curriculum Division, DLI.

APPLIED LANGUAGE LEARNING

PB 65 - 91 -1

1991 • Volume 2 • Number 1

EDITORIAL BOARD

Lidia Woytak, Editor	Curriculum Specialist, DLI
Kathleen M. Bailey	Head, TESOL/ESL Program, Monterey Institute of International Studies
David P. Benseler	Editor, *The Modern Language Journal*
John L. D. Clark	Dean, Program Evaluation, Research, and Testing, DLI
Ray T. Clifford	Provost, DLI
Annemarie Frazier, Word Processing Consultant	Academic Coordinator, School of Romance Languages, DLI
Martha Herzog	Dean, Curriculum and Instruction, DLI
Betty Lou Leaver	Dean, Slavic School, DLI
John A. Lett, Jr.	Director, Research and Evaluation, DLI
June K. Phillips	Director, Tennessee Foreign Language Institute
Gary D. Walter	Librarian, DLI

Linguistics & Language Behavior Abstracts

LLBA

Now entering our 24th year (135,000 abstracts to date) of service to linguists and language researchers worldwide. LLBA is available in print and also on-line from BRS and Dialog.

Linguistics & Language Behavior Abstracts

P.O. Box 22206
San Diego, CA 92192-0206
Phone (619) 695-8803 FAX (619) 695-0416

Fast, economical document delivery available.

APPLIED LANGUAGE LEARNING

PB 65 - 91 - 1

1991 • Volume 2 • Number 1

CONTENTS

1 Second-Language Learning in Adults: Correlates of Proficiency
Robert C. Gardner

29 Inside the Classroom: Learning Processes and Teaching Procedures
Leo van Lier

71 Measurement and Research Implications of Spolsky's Conditions for Second Language Learning
John L. D. Clark and Francis O'Mara

114 Information for Contributors

Applied Language Learning
Vol. 2, No. 1, pp. 1-28

Second-Language Learning in Adults: Correlates of Proficiency

Robert C. Gardner
University of Western Ontario

Are the problems facing adult second language learners really that different from those facing children, and are the underlying processes any different? There is a common belief that children are much more successful at language learning than adults, and there are many explanations for this presumed superiority. Foremost among these explanations is that of the critical period (Lenneberg, 1967) in which it is argued that the brain is relatively flexible until around the age of puberty, facilitating language acquisition before that time. Certainly most of us have known young children who seem to have mastered two languages with apparent ease, who function very comfortably in both of them, and who transfer from one to the other with very little hesitation. This is in stark contrast to many students in elementary, secondary, and/or post-secondary language courses who appear to find language learning very difficult, and who fumble to express even the simplest of ideas in the second language.

What research that has been done, however, generally questions the conventional wisdom that children are necessarily any more successful than adults. At least three reviews of the research literature (Genesee, 1978; Smythe, Stennett & Gardner, 1975; and Stern, 1976) have concluded that the bulk of the research literature favors the adults, even though the reviewers offer different possible interpretations for this. Smythe, Stennett and Gardner (1975) reviewed studies making direct comparisons between subjects at different ages, as well as research on Foreign Language in the Elementary Schools (FLES) programs in both the U.S.A. and Canada, and concluded that older individuals are more efficient learners than younger ones. They nonetheless favored the early introduction of second language training to young children for two reasons, (a) it permits more time for language learning to take place, and (b) young children are more likely not to have developed attitudinal/motivational characteristics that could impede successful language learning. Stern (1976) in his review focused attention on a ten year project in Great Britain that evaluated the teaching of French in elementary schools (Burstall, Jamieson, Cohen &

Hargreaves, 1974), as well as on immersion language programs in Canada. He argues that instead of focusing on an optimal age, we should instead search for those factors that promote second language acquisition at differing ages realizing that any particular age level has its own particular strengths and weaknesses. In his view, the immersion program has been one effective Canadian approach to this problem, but he stresses that the social context could influence the effectiveness of any program. Genesee (1978) similarly concludes that older students are more efficient language learners, though he recommends early language study because it provides more time for individuals to learn the language. He argues too that researchers and educators should concentrate on the development of pedagogical techniques that are appropriate to students at various ages.

At least one study has suggested another factor that should be considered when considering adult second-language learning. Snow and Hoefnagel-Höhle (1978) demonstrated that adults were more successful at language learning than children, but they also stressed that there are large individual differences in both children and adults in various aspects of second-language proficiency. This observation is particularly germane to the present discussion and serves as the basis for the rest of this review. That is, rather than being concerned about whether or not adults can learn a second language to any given degree of proficiency, it is recognized that among adults, as among children, there are varying degrees of success. Given this, the question arises as to whether or not there are any relatively consistent predictors of these differences in achievement.

When one directs attention toward adult language learning, there are a number of factors that one must consider. A very basic one has to do with the nature of the adults themselves. For the purposes of this review, we have taken adult to mean anyone who is university-or college-age or over. As a result, there is the possibility of great diversity from one study to the next in the nature of the sample, the type of program, the objectives of the students, the opportunities to use the language, etc. In addition, in many studies involving adult students, there is the possibility of considerable diversity within the sample itself. Even in an introductory university language course where one might expect more uniformity than in other adult language programs, there can still be considerable individual variation in reasons for being registered in the class, objectives, intent for further study, prior training in the language, etc. And such heterogeneity can complicate relationships of some variables with indices of proficiency. These differences are simply exacerbated in other adult language programs.

There are many factors that might influence second-language achievement, and thus there are many variables that might correlate with second-language proficiency. In discussing this problem, Taylor

(1974) argued that, in comparison with children, adults tend to be more developed conceptually, to have a more clearly articulated self-identity, and to have less time overall to devote to language study. Such factors could, he felt, influence second language achievement in adults; they also could interact with one another to affect what may or may not be potential correlates of second-language proficiency.

The conceptual development of adults could be both an asset and a drawback to language learning. Because of their conceptual development, adults have a considerable amount of world experience and own language skills that they bring with them to the learning environment. Such material can thus serve as conceptual pegs on which to hang new language material. Adults have problem-solving experience, and the skills they develop as a result can be used in the language learning situation, though they may or may not be facilitative. Adults have a well-developed level of intellectual functioning, and they are used to communicating at this level. The language learning context, however, often leaves them without the necessary linguistic code to communicate their thoughts, and they are forced to speak or write in a way that is considerably below their level. The majority of second-language learning situations also are relatively artificial, and adults' lack of experience with such artificiality can influence their level of language acquisition. Finally, simply as a function of their age, there can be considerable variation within a language classroom even in prior exposure to the second language or other languages. Such prior experiences can have an influence (cf. Eisenstein, 1980; Spurling & Ilyin, 1985) on second language acquisition, or even reactions to the context itself. When considering correlates of second language proficiency, therefore, the mere conceptual development of the adult, and individual differences in this development, should be considered when interpreting the data from any one study or when comparing the results from one study to another.

The well-articulated sense of self-identity and individual differences in this respect also can influence second language acquisition. For some individuals, the language they speak is an integral part of their own self-identity (cf. Giles & Byrne, 1982) and learning and using a second language represents an attack on their perceptions of self. Taylor, Meynard and Rheault (1977) demonstrated, in fact, that the fear that learning a second language might cause individuals to lose the link with their own culture was associated with low self-ratings of second language proficiency. In addition, attitudes toward the other language group have been shown in both children and adult samples to be related to proficiency in a second language (Gardner, 1985), thus it is not only own group loyalty, but also reactions toward the other community that can influence the development of second language proficiency in adults. Even in general terms, one's sense of well-being can be questioned,

regardless of ingroup/outgroup attitudes, by the potential embarassmemt that could arise if one fails to express oneself articulately, or is forced to use a form of language that one knows is inappropriate. Such considerations suggest that both attitudinal and personality variables could be implicated in second language acquisition in adults.

The time available to adults to learn a second language is another factor that could materially influence the degree of language acquisition. Generally, there are many pressures operating on an adult from his or her career, to family obligations, to social responsibilities, and these detract from the task of language learning. Thus, much proficiency must be acquired within given time constraints, and individual differences in characteristics that promote efficiency in learning (such as intelligence and language aptitude) would be expected to influence achievement. Finally, these outside pressures could be expected in some situations to distract some learners more than others, and the extent to which they do could be expected to vary as a function of the individual's specific purpose for language study and degree of motivation to persevere.

This overview was intended to focus on the potential factors operating on adult language learners and to indicate just how as a result any number of individual difference variables could correlate with proficiency in a second language. The following review will consider four classes of possible correlates, attitudes and motivation, language aptitude and intelligence, language learning strategies, and personality characteristics. As indicated above, there is good reason to assume that such variables, which have been shown to predict proficiency among school-age children (cf., Gardner, 1985, for reviews) should show similar if not higher relations among adults. As also indicated above, however, and which will be considered in more detail in what follows, the greater heterogeneity on all attributes of any particular group of adults studying a second language makes the investigation of such correlates particularly troublesome -- more so than with classes of school children who can be expected to have much more similar backgrounds.

Attitudes and Motivation

The impetus for research on the role of attitudes and motivation in second language study, much of which centers on children, has, as one of its foundations, observations made on adults. In an investigation of bilingual language behavior, Lambert (1955) noted one American graduate student who was dominant in French. Interviews with this student uncovered a strong identification with the European French and a disillusionment with the U.S. He had recently visited France and intended to return as soon as possible. Comparable observations were made by Whyte and Holmberg (1956) who found that Americans

workin[illegible] Latin America who identified with Latin Americans and sough[illegible] nteract with them socially were more successful at learning the language than those who didn't make such attachments. In a somewhat different context, Nida (1956) described the case of a foreign-born American missionary who was unable to learn a language that he needed for his missionary work. Nida claimed that the man had excellent teachers, a high level of intelligence, and a strong motivation to learn, and attributed his lack of success to his strong identification with the American culture. When he immigrated to the U.S. as a child, he worked very hard to learn English, insisted that English be spoken in the home, and dissociated himself from his ethnic background. Nida suggested that the man's strong identification with English prevented him from learning any other language.

This concept of identification underlies much of the research that has been conducted on the role of attitudes and motivation in second language acquisition. Lambert (1963) was the first to provide a theoretical model linking perceptions of self-identity, reactions to the other language group, and motivation to learn a second language with actual proficiency in that language, and the influence of that model is still very great (see also Lambert, 1967; Gardner & Lambert, 1972). Although much of the earlier research focussed on school children, it is very clear that Lambert's (1963; 1967) formulations apply to adults. The association is made even more explicit in Lambert's (1974) discussion of the role of such factors in second language acquisition when he introduced the distinction between additive and subtractive bilingualism.Although clearly applicable to young children, these potential consequences of bilingualism are most clearly applied to the adult community. Moreover, Lambert's (1974) discussion of the role of subtractive bilingualism in the immigrant community largely addresses the issue as it relates to the adult in that group. The process can take place in both children and adults, to be sure, but the long-term consequences are found in the adult.

A number of empirical investigations on the role of attitudes and motivation in second language learning have been conducted with adults, and, although the results are mixed, there nonetheless appears to be sufficient evidence to suggest that such variables are implicated. Some of these studies have been concerned with university students. Within this group, Mueller (1971) found that attitudes toward French people correlated positively with grades in French, while Mueller and Miller (1970) obtained similar correlations with students' reports of their grades. Taylor, Meynard and Rheault (1977), on the other hand, found that threat to ethnic identity resulting from studying English as a second language correlated negatively with self-ratings of proficiency among francophone university students.

Some studies have made use of modifications of the Attitude/Motivation Test Battery (AMTB) (Gardner, 1985) and have employed multivariate analyses. Gliksman (1981), for example, conducted two studies using factor analytic procedures. In the first study, he obtained one factor which he referred to as an *AMI* (Attitude/Motivation Index) factor, that shared variance in common with most of the attitudinal/motivational measures and indices of French proficiency. In the second study, a similar dimension, identified as an *Integrative Motive*, did not obtain high loadings from measures of French achievement--even though in both studies an AMI index formed by aggregating scores from the attitude and motivation scales correlated signficantly with French achievement and with various indices referring to behavioral intentions involving French. Lalonde and Gardner (1984) and Gardner, Lalonde and Pierson (1983) used causal modeling procedures (LISREL) to test specific models. Both of them provided clear support for the socio-educational model of second language learning (Gardner, 1985) showing causal links between attitudes, motivation and second language proficiency.

This is not to suggest that the socio-educational model is necessarily the only one that could be appropriate (cf., Gardner, 1985, p. 166). Other models have been proposed such as Clément's (1980) social context model, Giles and Byrne's (1982) intergroup model and Schumann's (1978) acculturation model, that include attitudinal/motivational variables as functional elements. Research with adult subjects attests to their validity (see, for example, Clément, 1987; Hall & Gudykunst, 1986; and Schumann, 1986). The models are each developed for somewhat different contexts and purposes, and emphasize somewhat different variables, but in each model attitudinal/motivational variables are implicated, and the relevant research generally demonstrates that they do have a direct or indirect link with second language proficiency.

Contrasting results were reported by Teitelbaum, Edwards and Hudson (1975) in their investigation of American university students learning Spanish. Their study involved two different samples, those claiming Spanish identity and those not, and considered the relationship of 109 predictors (mostly single items) to Spanish proficiency defined in terms of Cloze test scores. Very few of the relationships, at either the correlational or factor analytic level, were significant.

A number of other studies dealt with adults who were studying a second language, but not as part of a university program. Lambert, Gardner, Barik and Tunstall (1963) investigated adult students (primarily from the U.S.) in an intensive French language summer school program. Students were classified as either elementary or advanced based on an assessment of their French language skills. At the elementary level, French proficiency was associated with francophilia and an integrative orientation in language study. At the advanced

level, attitudes and orientation were not factorially related to proficiency. In fact, at the level of correlations, francophilia was positively associated with proficiency among elementary students but negatively among advanced ones. Gardner, Smythe and Clément (1979) conducted a similar study using the AMTB with samples of Canadian and American adults, also in an intensive summer school French program. Factor analyses conducted on each sample yielded three comparable factors identified as *Integrative Motive*, *French Proficiency*, and *Anxiety*. For the Canadian sample, but not for the American one, indices of French Oral Expression, satisfaction with outcomes and the quality of the program shared variance in common with the *Integrative Motive* factor.

Such inconsistent results were also obtained in other studies. Positive results are reported by Jacobsen and Imhoof (1974), who found that Japanophilia was among the best three predictors of Japanese speaking proficiency for both men and women missionaries in Japan, Spolsky (1969), who demonstrated that grades in English were associated with similarity of perceptions of self and English speakers among foreign students in the U.S., and by Oller, Hudson and Liu (1977) who found an association between ratings of Americans and proficiency in English among Chinese university students in the U.S. Negative results involving attitudes toward English speakers, however, were obtained by Chihara and Oller (1978) with Japanese adults studying English in Japan, while inconsistent results were reported by Oller, Baca and Vigil (1977) in their study of Spanish women enrolled in a Job Corps vocational school program. Inconsistent results were also reported by Ely (1986) who found that strength of motivation correlated significantly with only one of the six criteria investigated. Clearly, therefore, although there is some evidence that attitudinal/motivational variables relate to second language proficiency among adults, the results are far from consistent.

There are many possible reasons for the inconsistency in the results obtained. Different researchers use different measuring instruments and often different definitions of the various constructs, and the implications of these methodological issues have seldom been investigated, or even considered. Then, too, there are many other factors to be considered when conducting this type of research. Some of the more obvious ones are considered later in this article in the section on potential research problems.

Language Aptitude and Intelligence

There is a long history of research on language aptitude (see, for example, Henmon, 1929; Kaulfers, 1939; Carroll, 1974), but present day

conceptions of language aptitude are not much different than they were thirty years ago. For adults, the two major measures of language aptitude appear to be the Modern Language Aptitude Test (Carroll & Sapon, 1959) and the Language Aptitude Battery (Pimsleur, 1966), and the Defense Language Aptitude Battery (Petersen & Al-Haik, 1976).

The Modern Language Aptitude Test consists of five subtests, Number Learning, Phonetic Script, Spelling Clues, Words in Sentences, and Paired Associates, but is described (Carroll, 1974) as assessing four abilities necessary for the successful acquisition of a second language. These abilities are Phonetic Coding, Grammatical Sensitivity, Memory Ability, and an Inductive Language Learning Ability. They are presumed to be relatively enduring characteristics. In his survey of foreign language achievement of university and college language majors, Carroll (1967) reported that there was no evidence that prior language training influenced language aptitude, however, Carroll (1979) does point out that language learning abilities undoubtedly reflect environmental and educational influences. At least one study (Eisenstein, 1980) has shown, furthermore, that bilinguals scored higher on the Modern Language Aptitude Test as adults than do monolinguals.

The Language Aptitude Battery (Pimsleur, 1966) consists of six parts, two involving verbal reports of grade point average and ratings of interest respectively, and the last four comprising objective assessments. These latter four tests include measures of vocabulary, language analysis, sound discrimination, and a sound-symbol association ability. Conceptually these tests are similar to those in the Modern Language Apitude Test, and the presumed underlying abilities are also comparable. Pimsleur (1966) referred to a "talent" for languages, as distinct from abilities, and argued that his test measured the three primary components, Verbal Intelligence (including analytic reasoning), Motivation, and Auditory Ability.

The Defense Language Aptitude Battery (DLAB) (Petersen & Al-Haik, 1976) was developed to aid in the selection of candidates for the Defense Language Institute. It requires 80 minutes to administer, and provides scores on three factor scores and a total score. There are four parts to the test, and some parts have multiple sections. One part pairs pictures with an artificial language, and subjects are required to deduce appropriate language forms; another involves recognition of foreign language sounds; another requires subjects to form associations between sounds and symbols; and the last part involves the application of new grammatical rules in a translation task. All but the first part make use of aurally presented langage material. Not much validity information appears to have been published, but bivariate correlations involving total scores and multiple correlations, using the parts with course grades as criteria demonstrate that the Defense Language Aptitude

Battery performs comparably with both the Modern Language Aptitude Test and the Pimsleur Language Aptitude Battery.

The reliability and validity data for both the Modern Language Aptitude Test and the Language Aptitude Battery as indicated in their manuals and subsequent research articles are both quite impressive (see, for example, Carroll and Sapon, 1959; Pimsleur, 1966). Carroll (1962) shows, furthermore, that the correlation between scores on the Modern Language Aptitude Test and performance in second language classes tends to increase from high school to university to more intensive language courses. He argues that, where there is considerable opportunity to learn, individual differences in motivation, and/or variability in teaching effectiveness, etc., the aptitude-achievement association will be moderated. However, given a rigorous language program where these other variables do not have a chance to operate, he proposes that individual differences in aptitude will account for the most variance in second language proficiency. Moreover, his analysis tends to be supported by the data. Under conditions of intensive language training as in military contexts, scores on the Modern Language Aptitude Test are highly predictive of achievement. Although some work-sample tests of aptitude have been developed in the interim, the Modern Language Aptitude Test still appears to be a major selection test used by the military.

To a considerable extent, the rationale underlying the concept of language aptitude is that it represents a collection of abilities which, through a process of transfer, facilitate the learning of second language material. In the assessment of such aptitude, there is consequently an emphasis on vocabulary and grammatical skills in one's own language. In addition, because of the need to remember language material, often presented orally, the concept also includes the notion of a memory ability, the capacity to distinguish among sounds, and the facility to somehow code these sounds so that they can be retrieved later. Carroll (1979) has provided an overview of language aptitude as it relates to the broader area of mental abilities. By contrasting research on the MLAT with factor analytic investigations of mental abilities, Carroll demonstrates that language aptitude is much more than simply the verbal components of intelligence. There are correlations, to be sure, but it would obviously be too simplistic to equate language aptitude with verbal ability as it is conceived in factor analytic investigations of intelligence.

Relatively little recent research appears to have been conducted, particularly using adult students, that has as its focus the identification of different abilities that might be included in assessing language aptitude. Since there have been marked changes in the entire area of cognitive psychology, ability testing, and approaches to language teaching (Stern, 1983), it is conceivable that other abilities could be identified

that might improve prediction of subsequent achievement and help to clarify the processes involved. It is perhaps a tribute to the usefulness of the tests as they are now constituted that such research is not being conducted.

This is not meant to imply that the concept is considered as important by all researchers. It is an explicit component of Gardner's (1985) socio-educational model of second language learning, but it plays a somewhat secondary role in Krashen's (1981) Monitor Model of second language learning. In Krashen's model, language aptitude is not considered important to language acquisition, only to conscious language learning with its stress on classroom related mechanics and monitored language behavior. Since language acquisition involves the development of an unconscious understanding and appreciation of the language, it is seen as more valuable, hence, the somewhat secondary role of language aptitude. Neufeld (1974) is even more critical of the concept, questioning whether it has any meaning at all given that most individuals acquire a first language, at least. He argues that aptitude is important only for accounting for the learning of mere frills of the second language that are of interest primarily to language teachers, test constructors and the like. Although it is an interesting argument, it tends to ignore the great diversity even among native speakers in terms of vocabulary, fluency, grammatical use, etc., that would also covary with aptitude (cf. Carroll, 1979). At any rate, these are perhaps two reasons for the decreased research activity on language aptitude.

Another possible reason is that researchers seem to be directing their attention to other variables which have some similarities to the concept of language aptitude and yet are somewhat distinct. One such example is language learning strategies which developed out of a consideration of the "good language learner." and focuses on those behaviors that seem to distinguish good language learners from those who aren't so good. One interesting study in this tradition was conducted by Nation and McLaughlin (1986) who compared expert adult language learners (e.g., bilinguals and multilinguals) with monolinguals on a laboratory artificial language task. There were two versions, an implicit-learning task and an explicit-learning task, which differed in terms of the extent to which subjects were instructed to look for rules. The results demonstrated that multilingual subjects outperformed the other two groups on the implicit task, but that the three groups were comparable on the explicit one. The researchers proposed that, when left to their own devices, multilingual subjects made use of strategies to help them organize unfamiliar language behavior. Such results suggest that, depending upon the language learning environment, individuals might differ in the extent to which they use optimizing strategies. Such a possibility seems to be an important one to

keep in mind when considering the role of language learning strategies in second language learning.

Language Learning Strategies

The term "language learning strategies" is a broad one with far-reaching implications. Reiss (1985) defines a strategy as "a conscious approach used by an individual to facilitate learning" (p. 513), and by implication emphasizes that it is an individual difference characteristic of the learner. Paivio (1983), on the other hand, proposes that "a language learning strategy refers to the skillful planning and management of language learning as carried out by the learner or language teacher" (p. 189), thus implicating both the learner and the teacher, and implying that some strategies may be better than others. Other researchers (for example, Rubin, 1975; Stern, 1975) have proposed behaviors and/or strategies that differentiate between successful and non-successful language learners. These include such things as a willingness to guess and a propensity to be correct, a desire to communicate, active participation in language learning tasks, a search for meaning, monitoring both own and others' language, and experimenting with different forms, etc.

It is obvious from such a list that there are many different types of strategies that could be proposed to improve second language proficiency. Some of them are definitely learner strategies in that they are activities used by the individual student to promote proficiency. Others, however, seem to be more closely linked with teaching strategies in that they are procedures used to promote instruction. An example would be programs that use pictures to facilitate vocabulary acquisition, or those that use pattern drills, etc., to promote fluent discourse. Although students might incorporate them into their own repertoire, thus making them learner strategies, until this happens, it seems better to characterize such procedures as teaching strategies.

Still other strategies seem to focus more on communication than learning. Admittedly students learn language material through communication, but many of these strategies (e.g., Stern's, 1975 'willingness to use the language in real communication') seem better characterized as a communication strategy. Finally, some strategies seem more to reflect personality attributes. Being extroverted, uninhibited, prepared to take risks, etc., may (or may not) facilitate language learning, but if they too are considered as strategies, the term becomes all too inclusive. Similar observations have been made by others, of course. Bialystok (1984), for example, discusses the importance played by learning strategies, communication strategies and social strategies.

Similarly, Reiss (1985) argues that a number of the strategies or styles in Rubin's and Stern's list should be differently classified.

Even if one focuses attention on strategies as determined by the individual, there remains some degree of ambiguity, and their measurement is not straightforward. Reiss (1985), for example, investigated seven learning strategies, guessing, motivation to communicate, attending to form, practicing, monitoring, attending to meaning, and using mnemonics, and developed items to measure each of these strategies. Some of these items, however, seem to include a motivational component as well. Thus, the item:

> If you have an opportunity to practice the foreign language you are studying with a native speaker do you:
>
> (a) plunge right in, even though you may appear foolish
>
> (b) only speak the foreign language with the person after you have known him/her for some time
>
> (c) speak only English with him/her

is conceptually very similar to items included to measure Motivational Intensity on the Attitude/Motivation Test Battery (Gardner & Smythe, 1981).

There have been a few attempts at categorizing language learning strategies, but the most comprehensive appears to be that suggested by Oxford-Carpenter (undated). She refers to distinctions made by other writers between primary vs. support strategies, direct vs. indirect, cognitive vs. metacognitive, etc., then provides a Taxonomy of Second Language Learning Strategies. This taxonomy distinguishes between Direct or Primary Strategies and Indirect or Support Strategies. Under the former she includes such specific techniques as Inferencing, L1-L2 (Contrastive), Mnemonic, General Memory, Emphasis and Summarizing, Functional Practice, Formal Practice with Sounds or Rules, Clarification and Verification, and Resourcing. Instances of indirect strategies include Attention-Enhancing, Self-Management, Affective, Planning and Goal Setting, Production or Reception, Study, Practice-Creating, and Cooperation strategies. This taxonomy is ripe with examples and definitions under each category. Oxford (1986) described the development of the Strategy Inventory for Language Learning (SILL) which is a self-report assessment of the use of such strategies. This inventory appears to provide the most complete assessment of language learning strategies to date, and describes research that suggests its potential value.

Research that focuses on individual differences in language learning strategies has often made use of interviews or student diaries to

identify the strategies used (see, for example, Bailey, 1983; Naiman, Fröhlich, Stern & Todesco, 1978) as well as questionnaires such as those used by Reiss (1985), Oxford, Nyikos & Crookall (1987), and Politzer and McGroarty (1985). Often the results are largely statements concerning the strategies used (Bailey, 1983; Naiman et al., 1978) by successful students. At other times individual differences in various strategies are related to self-ratings of proficiency, possible personality variables, motivational variables and the like. Research by Politzer (1983) and Politzer and McGroarty (1985) has demonstrated that factors such as motivation, language proficiency, and the purpose for studying a given language, etc., are related to which strategies are used.

The most detailed investigations of language learning strategies have been conducted, however, by Oxford, Nyikos and Crookall (1987) who factor analyzed responses to the Strategy Inventory for Language Learners made by young adults in a military language program. The five major factors obtained were *General Study Habits*, *Functional Practice*, *Speaking and Communicating Meaning*, *Studying or Practicing Independently* and *Mnemonic Devices*. These results suggest that there are relatively independent strategies involved in language learning, and it would seem important to direct further research to determining factors that influence the use of these strategies, and the effect that the use of these strategies has on language proficiency.

Experimental studies of language learning (or perhaps, more accurately, teaching) strategies have tended to focus on mnemonic techniques that involve imagery (i.e., nonverbal) procedures and have indicated their superiority in some aspects of language learning. The most popularly investigated procedures are the Keyword method (Atkinson & Raugh, 1975), and the Hook technique (Paivio, 1978). The studies involve contrasting the relative degree of success on specific language learning tasks of students instructed in the use of these procedures, as contrasted with students using other methods (e.g., synonym generation) or their own (Control) devices. As such, they are more like teaching strategies rather than learner strategies, but to the extent that the individual ultimately adapts them to their own purposes, they could be considered learner strategies. Despite their obvious superiority in the laboratory-based experiments (see, for example, Paivio & Desrochers, 1981; Pressley, Levin & Delaney, 1982), the use of mnemonic language learning strategies is not common in inventories of language learning strategies. Reiss (1985) reports that mnemonic strategies had the lowest frequency of occurrence of the seven she investigated among 98 college level students of a second language.

As indicated by this brief review, the concept of language learning strategies could be particularly useful in the study of adult language learning. Because of their experience with many problem-solving situations, adults undoubtedly approach second language learning with

various types of strategies which may or may not be productive. Research that helps to identify the more beneficial strategies, or which determines consistent correlates of such strategies, could be particularly useful in explaining the language learning process. It seems clear, however, that the nature, identification and measurement of such strategies require careful scrutiny, otherwise the concept itself will become too diffuse, tending to incorporate most variables associated with language learning.

Personality Variables

In discussing the role of personality in second language acquisition, Brown (1973) differentiated between two classes of personality variables, egocentric factors and social variables. Instances of the former included imitation, egoism and inhibition, while examples of the latter were empathy, introversion/extroversion and aggression. He also suggested a possible third class of personality-type variables which he referred to as cognitive style. Subsequent research has centered on various personality variables, many of which can be subsumed under those general categories, and in general, although the results are suggestive, they are by no means definitive.

Relatively few studies appear to have investigated a number of personality variables simultaneously, and their relation to proficiency in a second language, even though such studies would seem to be basic. Two studies were conducted in Sweden by Oskarsson (1975), one involving farm management students who were studying English as a second language, and the other involving vocational counsellors. Of the eight personality variables investigated in the first study, only one, "deliberateness" correlated significantly (and negatively) with English achievement. In the second study, measures of English reading proficiency correlated significantly positively with two personality variables, sociability and emotionality, and significantly negatively with deliberateness.

A study conducted by Lalonde (1982) used factor analysis to investigate the relationships among 19 personality variables, attitudinal/motivational variables, language aptitude and second language achievement among English speaking university students studying French. The results demonstrated that personality variables shared very little variance in common with measures of French proficiency. Four factors were defined primarily in terms of personality variables, and these were defined as *Conservatism, Affect, Sociability, and Analytic Orientation/Integrativeness*. This latter factor demonstrated a relationship between personality variables reflecting an analytic orientation similar to that discussed by Krashen (1981) and attitudinal/

motivational variables reflecting integrativeness. Such an association had been postulated by Krashen (1981).

Some studies have focused attention on one or two personality variables and their relation to second language proficiency. In these instances, the major personality variables appear to be Sociability (or related concepts), *Risk-Taking*, *Empathy*, *Field Dependence/Independence* and *Anxiety*. As is true of affective variables, in many cases much of this research has also involved children, but the following discussion considers only those studies conducted with adults.

Since language acquisition obviously involves interaction with other individuals, it is quite reasonable to expect a dimension like sociability or extraversion to relate to second language proficiency. Although sociability was not related to second language proficiency in Lalonde's (1982) study, evidence for a relationship is provided by Chastain (1975) also studying university level students. He reports significant positive correlations between sociability and grades in German and Spanish, but not French. A study using university students, conducted by Smart, Elton and Burnett (1970), on the other hand, suggests that sociability is negatively related to second language proficiency once ability is controlled, while a study by Ely (1986) found no relationship between Language Class Sociability and language proficiency. Such results are at best equivocal, but then so are expectations. Naiman, Fröhlich and Stern (1975) found, for example, that some teachers felt that good language learners were extraverted, while others felt that they were introverted. Kawczynski (1951), on the other hand, proposed that both introversion and extraversion facilitate language learning, depending upon the nature of the program. Smart et al. (1970) make a similar suggestion with respect to the role of sociability.

Discussions of the good language learner often refer to the willingness of such individuals to take chances, be good guessers, and make active use of the language (cf. Rubin, 1975; Stern, 1975). In discussing this material, Beebe (1983) argues persuasively that such strategies reflect a personality dimension of risk-taking, and goes on to link this concept to perceptions of self-identity, feelings of self-esteem and motivational determinants of second language proficiency. The analysis seems particularly appropriate to adult populations and worthy of further empirical investigation; at the present time, however, there does not appear to be any data directly linking individual differences in risk-taking as a personality attribute to proficiency in adult second language skills.

Field dependence/independence is another variable that many researchers consider as a possible correlate of second language achievement. Schumann (1986) lists it as a cognitive factor, while Brown (1973) views it as a merger of cognitive and affective variables. The variable itself is clearly quite complex and seems meaningfully considered as a

personality variable. Witkin, Goodenough and Oltman (1979) propose that field dependent individuals are sensitive, interested in others, and give the impression of being outgoing, while field independent individuals tend to be self-sufficient, analytical and appear to be somewhat aloof. Naiman, Fröhlich and Stern (1978) suggest that field independent individuals would be successful second language learners because they could separate language material to be learned from the larger context. Some support for this hypothesis appears to be provided by a study by Hansen and Stansfield (1981) who investigated university students of Spanish. The results demonstrated significant relations between field independence and Spanish achievement. The results were somewhat equivocal, however, because many of these relationships disappeared once the effects of scholastic ability were removed.

The concept of anxiety has also been linked to second language achievement. In one of the earlier studies, Chastain (1975) obtained negative relations between a test anxiety score and grades in Spanish among university students, but similar results were not obtained with students in French or German, or with other measures of anxiety. In discussing these results and others, Scovel (1978) proposed that facilitating anxiety could promote second language acquisition while debilitating anxiety could inhibit it, thus permitting both positive and negative correlations between anxiety and second language proficiency.

In later studies, many of the correlations are negative, suggesting that anxiety has a debilitating effect. Gardner (1985) summarizes many studies dealing with measures of *French Use* and *French Classroom Anxiety*, some of which were conducted on university level students, and argues that the results suggest that anxiety has a deleterious effect on proficiency. A similar conclusion was drawn by Trylong (1987) who studied the relation of French classroom anxiety to proficiency among university students registered in French. She found negative correlations with achievement on oral quizzes, written exams and coursegrades in a total sample of 216 students. Similar evidence of a negative relationship between anxiety and second language proficiency is reported by Clément (1987) with Francophone university students. He shows that proficiency in oral and written English is related to *Self-Confidence* with English in both majority and minority group contexts. In Clément's (1987) research, however, *Self-Confidence* is an aggregate measure comprised of high self-ratings of English proficiency, and low scores on English Use Anxiety.

Clearly, there are many aspects of language anxiety that could account for negative relations with proficiency. Some components other than self-confidence are suggested by Horwitz, Horwitz and Cope (1986) who have developed a Foreign Language Classroom Anxiety Scale. According to them, such anxiety has three components, Communication Apprehension, Test Anxiety, and Fear of Negative Evaluation in the

Foreign Language Classroom. Reliability and validity data for their scale are generally impressive (also see Horwitz, 1986). Scores on this test tend not to correlate very highly with indicators of general trait anxiety, but have moderate correlations with indices such as test anxiety. Moreover, scores on this scale tend to correlate negatively with grades in the second language.

Obviously, more research is required, but as this review suggests, there is evidence indicating that a number of personality variables correlate meaningfully with measures of second language proficiency. To date, insufficient attention seems to have been directed toward adult populations, and there is still no comprehensive model that provides a process explanation as to why specific personality variables should or should not facilitate second language acquisition. It would seem fruitful to direct attention toward multivariate studies of many potential personality correlates in order to determine whether there are any common underlying dimensions associated with language proficiency. Such multivariate approaches have been employed with language aptitude, and with attitudinal/motivational variables with some degree of success, and would seem highly applicable with respect to personality correlates.

Potential Research Problems

As the previous review has indicated, there is evidence to conclude that, among adult samples, there are relations between indices of proficiency in a second language and measures of attitudes and motivation, language aptitude and intelligence, language learning strategies, and some personality variables. The results, however, are not unequivocal, nor are they necessarily unconfounded. The purpose of this section is to highlight potential problems in conducting such research which are exacerbated when dealing with adult samples and which might have, or could influence any results obtained. These problems include the heterogeneity of adult samples with respect to prior background, education and language learning experiences, possible contextual effects produced by the language learning situation, sampling distributions of statistics which are reflected in Type I and Type II errors, the quality of the measuring instruments for both the correlates of language proficiency and the proficiency itself, and possible confounds of proficiency with the potential correlating variable.

A major problem that could obfuscate conclusions drawn from adult samples is the very real heterogeneity among the adults themselves on variables that might influence second language acquisition. Let us assume, for example, that language aptitude, motivation and previous experience with a second language all do in fact influence how

well an adult will perform on some measure of second language proficiency. Assume, further, that in the population at large these three attributes are relatively independent. Selection factors associated with any particular sample could, however, artificially attenuate or exaggerate relations between these attributes and second language achievement. Among some adult samples there might be prior screening for aptitude, which would decrease the correlation of aptitude and proficiency. On the other hand, a language program might attract only those who are highly motivated and the resulting lack of variability in motivation would attenuate its correlation with proficiency.

It is also possible that confounding variables could influence the relationships obtained. Consider, for example, an adult language class in introductory Spanish offered in a university night class program. In such a class, the students could differ in aptitude, motivation, and prior experience with Spanish. It is meaningful to assume that those with prior training in Spanish would score higher on a measure of Spanish achievement than those lacking the training. If at the same time, however, those with the prior experience were lower on aptitude or motivation, this would attenuate the correlation between aptitude or motivation and achievement. That is, prior experience which could be due to any number of factors could influence the correlations obtained.

This is not meant to suggest that the above relationships would necessarily operate, or that these are the only variables that might relate to proficiency. The point is simply that adult language classes do not generally represent random samplings from the population. For any given class, there may be any number of reasons why the various students are there, and thus relationships of variables to proficiency can be influenced.

Often it is also the case that researchers interested in correlates of proficiency conduct their research in a number of different classes, with different teachers, sometimes with different objectives and/or pedagogical procedures, and sometimes even with students at different levels–introductory, intermediate, advanced, etc. If researchers simply correlate the measures of the characteristics of interest (e.g., aptitude, motivation, personality, etc.) with proficiency without somehow removing the potentially confounding effects of these class differences, the results could be quite misleading. Such class effects could be removed by computing within class correlations, or by first standard-izing the data within class before computing the correlations. Such pro-cedures then indicate how individual differences within class relate to proficiency, and thus eliminate effects due to class that could either artificially increase or decrease the apparent correlation. A hypothe-tical example of how correlations might be attenuated because of differences due to class related attributes is presented by Gardner (1985).

A second problem that merits consideration is the sampling dis-

tribution of a statistic–often in this type of research the Pearson product-moment correlation coefficient. When a researcher concludes that a given correlation is significant at say the .05 level, the implication is that, if the true correlation were 0, one would expect a value as large as that obtained less than 5% of the time on the basis of chance. That is, the correlation coefficient has a sampling distribution which is a function of the value of the correlation in the population (i.e., the true correlation) and the sample size. A correlation of .36 based on a sample size of 30 individuals is considered significant at the .05 level using a two-tailed test. That is, if the true correlation in the populatioon were 0, 95% of the time we would expect to obtain a correlation somewhere between -.36 and +.36, and only 5% of the time would we expect a value more extreme than ±.36. The type I error rate is said to be 5% in this case. If the sample size were 100 on the other hand, we would require a correlation.of only .20 to reject the null hypothesis and conclude that the population correlation was not 0 with a Type I error rate of .05. Not-withstanding this, the best point estimate we could make of the true correlation is the value we obtained in our sample.

But what if we knew that the correlation in the population was .30. Here again, we would expect sampling variation. Using the Fisher Z transformation, it could be shown that, with a sample size of 30, we would expect 95% of the correlations to range between -.07 and .60. More importantly, we would expect only 37% of them to be .36 or greater. Thus, with a sample size of 30, the probability of our rejecting the null hypothesis when it is false (and the true population value is .30) is only .37. For a sample size of 100 and a population correlation of .30, however, the 95% confidence interval ranges from .11 to .47. More to the point, the probability of our rejecting the null hypothesis when it is false (and the true population value is .30) is .86. That is, we have much greater power with a larger sample size.

In general, it seems meaningful to assume that the true correlation between different classes of variables such as motivation and achievement, or aptitude and achievement, or personality and achievement, would be around .30 to .40. And, if this is the case, the moral should be clear, viz., small sample sizes may lack the power more times than not to detect the relationship.

A third potential problem lies in the quality of the measuring instruments used to assess the possible correlates or the measures of proficiency. When developing any measuring instrument, it is important to obtain valid and reliable assessments of the constructs of interest. Validity refers to the extent to which the measure assesses what it is purported to measure, and is obviously important to permit meaningful generalization of results. Reliability refers to the absence of error in the measurement, and is important to provide stable assesments of the construct of interest. Obviously many factors can influence

reliability, but one that has a very direct effect is the number of observations one makes of the construct of interest. In individual difference research, this typically involves the number of items comprising the measuring instrument. That is, other things being equal, a ten item attitude scale will be more reliable than a five item scale. And, by and large, single item scales can often be quite unreliable. Such considerations are important in any investigation of individual differences, but they deserve mention in the present context because often when studying adults, time is at a premium, and a researcher will be tempted to shorten the scales used. If this is done, it is possible that the resulting scales will lack reliability (if not validity) with the result that relatively few significant correlations will be obtained.

Although these comments have been made in the context of the correlates of second language proficiency, they are equally appropriate to indices of proficiency themselves. Attention should also be directed toward reliable and valid measures of proficiency. Years ago many researchers employed self-ratings as one index of second language proficiency, and these took the form of ratings on seven point scales of reading, speaking, writing and understanding skills (see, for example, Gardner & Lambert, 1972). Often, too, these scales did not correlate very highly with other variables. More recently, Clark (1981) proposed can-do self-rating scales where individuals rate their skill on specific tasks in the language such as introducing oneself in social gatherings, understanding radio news broadcasts, etc. The resulting self-ratings of speaking skills, etc., obtained by aggregating responses on all the speaking tasks provide very reliable measures that correlate substantially with corresponding assessments of proficiency (Clark, 1981). By making the judgments more behaviorally specific and aggregating over a series of items, a much more useful measuring device was developed.

A somewhat related issue has to do with the possibility of confounds between a possible predictor of second language proficiency and the proficiency itself. Consider, for example, a possible language learning strategy item such as "I do well in languages because I use mnemonic techniques to help me remember vocabulary." On the surface, at least, this sounds like an item that assesses the use of a mnemonic language learning strategy. Obviously, agreement with such an item indicates the use of a mnemonic device, while disagreement indicates the lack of use. However, agreement with the item also indicates that the individual "does well in languages," while disagreement reflects poor performance. Since such perceptions tend to correlate to some extent with actual proficiency, a correlation of such an item with proficiency is ambiguous. A significant positive correlation

could indicate an association between the use of mnemonics and proficiency, perceptions of achievement and proficiency, or a combination of both. That is, there is a possibility of a confound.

Consider another example of an item written to assess language classroom anxiety. Subjects reacting to an item such as "I sometimes get so flustered in language class that I remember very little of what was taught" are indicating something about their anxiety in the classroom, but they are also indicating that they do not do well in the subject matter. A negative correlation between such an item and proficiency could therefore reflect such confounding. If the sole purpose of such measures is to predict proficiency, such confounds might be acceptable, but if they are used to explain processes underlying language acquisition they could lead to misinterpretations.

The preceding examples are somewhat exaggerated and were made up rather than taken from any particular scales, but they indicate possible problems that can arise. Particularly when correlations between one class of variables such as strategies, anxiety, attitudes, etc., and another such as language proficiency are overly high (e.g., greater than .3 or .4 in samples of about 100 individuals), it may well do to inspect the items for such possible confounds. After all, given the number of variables operating on an individual when acquiring a second language, it is unreasonable to expect any one class of variable to correlate too highly with another class -- and a reasonable estimate of high in this context is around .30. This may represent only 9% of the common variance, but when you think about it that is quite subtantial.

Conclusions

The purpose of this article was to review some research conducted on adult populations dealing with possible correlates or predictors of proficiency in a second language. These correlates refer to individual difference characteristics which in one way or another mediate or moderate the effects of language input in any language learning context. For some characteristics, there are theoretical models proposed to explain the way in which the individaual differences interface with language learning and actually become part of the process itself, while for other characteristics there does not appear to be much in the way of an underlying theoretical foundation.

In this review, I have considered four classes of individual difference variables, attitudes and motivation, language aptitude and intelligence, language learning strategies, and personality variables. In each case, there was evidence to suggest relationships with second language proficiency. There were also, however, a number of problems

evident in this research dealing with conceptual, methodological and analytic issues. As discussed, there are a number of general problems that one should be on the lookout for in reading about or conducting this type of research regardless of the population under consideration. However, these are even added potential hazards when the research is conducted on adult populations.

Obviously there are many potential problems that could be mentioned, but, in this presentation, I have focussed on five issues that could seriously influence any results obtained or generalizations made. Many of these are apparent or at least implicated in some of the published research, making clear generalizations and conclusions difficult. The problems discussed include the possible heterogeneity of adult samples, the effects that course-related charcteristics might have on potential correlations, the sampling distributions of statistics, particularly the correlation coefficient, the reliability and validity of measuring instruments, and the possibility of confounds of potential predictor variables with actual proficiency.

It seems clear that there is a growing activity in research with adults focussing on potential correlates of second langage proficiency. Such research can have important implications for selecting adults for various career-related language programs, and also for possibly identifying factors that should be considered in developing either career-related or recreationally-related language programs. The variables being investigated, however, although shown to have influential effects, are variables that play only a part in the language learning process. Their effects can be enhanced or muted, and researchers working in this area are advised to consider very carefully all the possible extraneous factors that can also play a role. The area itself is interesting, and the possible gains are many, but the relationships are delicate and require care in their investigation.

References

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of Russian vocabulary, *Journal of Experimental Psychology: Human Learning and Memory, 104*, 126-133.

Bailey, K. M. (1983). Competitiveness and anxiety in adult second language learning: Looking *at* and *through* the Diary Studies. In Seliger, H. W., & Long, M. H. (Eds.), *Classroom-oriented research in second language acquisition* (pp. 67-103). Rowley, MA: Newbury House.

Beebe, L. M. (1983). Risk taking and the language learner. In Seliger, H. W., & Long, M. H. (Eds.), *Classroom-oriented*

research in second language acquisition. (pp. 39-66). Rowley, MA: Newbury House.

Bialystok, E. (1984). Strategies in interlanguage learning and performance. In Davies, A., Cuper, C., & Howett, A. P. R. (Eds.), *Interlanguage* (pp. 37-48). Edinburgh University Press.

Brown, H. D. (1973). Affective variables in second language acquisition. *Language Learning, 23*, 231-244.

Burstall, C., Jamieson, M., Cohen, S., & Hargreaves, M. (1974). *Primary French in the balance*. Windsor, Berks.: NFER.

Carroll, J. B. (1962). The prediction of success in intensive language training. In Glaser, R. (Ed.) *Training research and education*. Pittsburgh: Pittsburgh Press.

Carroll, J. B. (1967). The foreign language attainments of language majors in the senior year: A survey conducted in U.S. colleges and universities. Cambridge, MA.: Harvard University Press.

Carroll, J. B. (1974). Aptitude in second language learning. Proceed*ings of the Fifth Symposium of the Canadian Association of Applied Linguistics*, 8-23.

Carroll, J. B. (1979). Psychometric approaches to the study of language abilities. In Fillmore, C. J., Kempler, D. & Wang, W. S-Y. (Eds.), *Individual differences in language ability and language behavior*. New York: Academic Press.

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test (MLAT)*. New York: Psychological Corporation.

Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning, 25*, 153-161.

Chihara, T. & Oller, J. W. (1978). Attitudes and attained proficiency in EFL: A sociolinguistic study of adult Japanese speakers. *Language Learning, 28*, 55-68.

Clark, J. L. D. (1981). Language. In Barrows, T. S. (Ed.), *A survey of global understanding: Final Report*, Chapter 9. New Rochelle, NY : Change Magazine Press.

Clément, R. (1980). Ethnicity, contact and communicative competence in a second language. In Giles, H., Robinson, W. P., and Smith, P. M. (Eds.) *Language: Social psychological perspectives*. Oxford: Pergamon Press.

Clément, R. (1987). Second language proficiency and acculturation: An investigation of the effects of language status and individual characteristics. *Journal of Language and Social Psychology, 5*, 271-290.

Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. *International Review of Applied Psychology, 29*, 159-172.

Ely, C. M. (1986). An analysis of discomfort, risk taking, sociability and motivation in the L2 classroom. *Language Learning, 36*, 1-25.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold Publishers.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA.: Newbury House.

Gardner, R. C., Lalonde, R., & Pierson, R. (1983). The socioeducational model of second language acquisition: An investigation using LISREL causal modeling. *Journal of Language and Social Psychology, 2*, 1-15.

Gardner, R. C., & Smythe, P. C. (1981). On the development of the Attitude/Motivation Test Battery. *Canadian Modern Language Review, 37*, 510-525.

Gardner, R. C., Smythe, P. C., & Clément, R. (1979). Intensive second-language study in a bicultural milieu: An investigation of attitudes, motivation and language proficiency. *Language Learning, 29*, 305-320.

Genesee, F. (1978). Is there an optimal age for starting second language instruction? *McGill Journal of Education, 13*, 145-154.

Giles, H., & Byrne, J. L. (1982). An intergroup approach to second language acquisition. *Journal of Multilingual and Multicultural Development, 1*, 17-40.

Gliksman, L. (1981). Improving the prediction of behaviours associated with second language acquisition. Unpublished Ph.D. dissertation, University of Western Ontario.

Hall, B. J., & Gudykunst, W. B. (1986). The intergroup theory of second language ability. *Journal of Language and Social Psychology, 5*, 291-302.

Hansen, J., & Stansfield, C. (1981). The relationship of field dependent-independent cognitive styles to foreign language achievement. *Language Learning, 31*, 349-367.

Henmon, V. A. C. (1929). Prognosis tests in the modern foreign languages. In Henmon, V. A. C. (Ed.). *Prognosis tests in the modern foreign languages*. New York: MacMillan.

Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly, 20*, 559-562.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*, 125-132.

Jacobsen, M., & Imhoof, M. (1974). Predicting success in learning a second language. *Modern Language Journal, 58*, 329-336.

Kaulfers, W. V. (1939). Prognosis and its alternatives in relation to the guidance of students. *German Quarterly, 12*, 81-84.

Kawczynski, A. S. (1951). The two psychological types of language students. *Modern Language Journal, 35*, 113-118.

Krashen, S. (1981). *Second language acquisition and second language*

learning. Toronto: Pergamon Press.

Lalonde, R. N. (1982). *Second language acquisition: A causal analysis*. Unpublished Master's thesis, University of Western Ontario, London, Canada.

Lalonde, R. N., & Gardner, R. C. (1984). Investigating a causal model of second language acquisition: Where does personality fit? *Canadian Journal of Behavioural Science, 16*, 224-237.

Lambert, W. E. (1955). Measurement of the linguistic dominance of bilinguals. *Journal of Abnormal and Social Psychology, 50*, 197-200.

Lambert, W. E. (1963). Psychological approaches to the study of language. Part II: On second language learning and bilingualism. *Modern Language Journal, 14*, 114-121.

Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social Issues, 23*, 91-109.

Lambert, W. E. (1974). Culture and language as factors in learning and education. In Aboud, F. E., and Meade, R. D. (Eds.) *Cultural factors in learning and education*. Bellingham, Washington: Fifth Western Washington Symposium on Learning.

Lambert, W. E., Gardner, R. C., Barik, H. C., & Tunstall, K. (1963). Attitudinal and cognitive aspects of intensive study of a second language. *Journal of Abnormal and Social Psychology, 66*, 358-368.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

Mueller, T. H. (1971). Student attitudes in the basic French courses at the University of Kentucky. *Modern Language Journal, 55*, 290-298.

Mueller, T. H., & Miller, R. I. (1970). A study of student attitudes and motivation in a collegiate French course using programmed language instruction. *International Review of Applied Linguistics, 8*, 297-320.

Naiman, N., Fröhlich, M., & Stern, H. H. (1975). *The good language learner*. Toronto: Ontario Institute for Studies in Education.

Naiman, N., Fröhlich, M., Stern, H. H., & Todesco, A. (1978). *The good language learner*. Research in Education Series No. 7, Ontario Institute for Studies in Education, Toronto.

Nation, R., & McLaughlin, B. (1986). Novices and experts: An information processing approach to the "good language learner" problem. *Applied Psycholinguistics, 7*, 41-56.

Neufeld, G. G. (1974). A theoretical perspective on the nature of linguistic aptitude. *Proceedings of the Fifth Symposium of the Canadian Association of Applied Linguistics*, 74-92.

Nida, E. A. (1956). Motivation in second-language learning.

Language Learning, 7, 11-16.
Oller, J. W., Baca, L., & Vigil, F. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of Mexican-Americans in the Southwest. *TESOL Quarterly, 11*, 173-182.
Oller, J. W., Hudson, A., & Liu, P. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of native speakers of Chinese in the United States. *Language Learning, 27*, 1-27.
Oskarsson, M. (1975). The relationship between foreign language proficiency and various psychological variables. Sweden: Gothenberg University, Language Teaching Research Center.
Oxford, R. L. (1986). Development of a new survey and taxonomy for second language learning. Paper presented at the Learning Strategy Symposium, New York.
Oxford, R. L., Nyikos, M., & Crookall, D. (1987). Learning strategies of university foreign language students: A large-scale factor analytic study. Unpublished manuscript.
Oxford-Carpenter, R. L. (undated). A new taxonomy of second language learning strategies. Washington, D.C.: ERIC Clearinghouse on Language and Linguistics.
Paivio, A. (1978). On exploring visual knowledge. In Randhawa, B. S. & Coffman, W. E. (Eds.), *Visual learning, thinking, and communication*. New York: Academic Press.
Paivio, A. (1983). Strategies in language learning. In Pressley, M. & Levin, J. R. (Eds.), *Cognitive, strategy research: Psychological Foundations*. New York: Springer-Verlag.
Paivio, A. & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology, 73*, 780-795.
Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery*. New York: Harcourt Brace Jovanovitch.
Politzer, R. L. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievement. *Studies in Second Language Acquisition, 6*, 54-68.
Politzer, R. L., & McGroarty, M. (1985). An exploratory study of learning behaviors and their relationship to gains in linguistic and communicative competence. *TESOL Quarterly, 19*, 103-123.
Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research, 52*, 61-91.
Reiss M-A (1985). The good language learner: Another look. *Canadian Modern Language Review, 41*, 511-523.
Rubin, J. (1975). What the good language learner can teach us. *TESOL Quarterly, 9*, 41-51.
Schumann, J. H. (1978). The acculturation model for second-language

acquisition. In Gingras, R. G. (Ed.), *Second-language acquisition and foreign language teaching*. Arlington, VA: Center for Applied Linguistics.

Schumann, J. H. (1986). Research on the Acculturation Model for second language acquisition. *Journal of Multilingual and Multicultural Development, 7*, 379-392.

Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning, 28*, 129-142.

Smart, J. C., Elton, C. F., & Burnett, C. W. (1970). Underachievers and overachievers in intermediate French. *Modern Language Journal, 54*, 415-422.

Smythe, P. C., Stennett, R. G., & Gardner, R. C. (1975). The best age for beginning foreign language training: Issues, opinions and facts. *Canadian Modern Language Review, 32*, 10-23.

Snow, C. E., & Hoefnagel-Hohle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development, 49*,1114-1128.

Spolsky, B. (1969). Attitudinal aspects of second language learning. *Language Learning, 19*, 271-283.

Spurling, S., & Ilyin, D. (1985). The impact of learner variables on language test performance. *TESOL Quarterly, 19*, 283-301.

Stern, H. H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review, 31*, 304-318.

Stern, H. H. (1976) Optimal age: Myth or reality. *Canadian Modern Language Review, 32*, 283-294.

Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford University Press.

Taylor, B. P. (1974). Toward a theory of language acquisition. *Language Learning, 24*, 23-35.

Taylor, D. M., Meynard, R., & Rheault, E. (1977). Threat to ethnic identity and second-language learning. In Giles, H. (Ed.), *Language, ethnicity and intergroup rrelations*. New York: Academic Press.

Teitelbaum, H., Edwards, A., & Hudson, A. (1975). Ethnic attitudes and the acquisition of Spanish as a second language. *Language Learning, 25*, 255-266.

Trylong, V. L. (1987). *Aptitude, attitudes, and anxiety: A study of their relationships to achievement in the foreign language classroom*. Unpublished doctoral dissertation, Purdue University.

Whyte, W. F., & Holmberg, A. R. (1956). Human problems of U.S. enterprise in Latin America. *Human Organization, 15*, 1-40.

Witkin, H., Goodenough, D., & Oltman, P. (1979). Psychological differentiation: Current status. *Journal of Personality and Social Psychology, 37*, 1127-1145.

Author

ROBERT C. GARDNER, Pofessor, Department of Psychology, The University of Western Ontario, London, Canada N6A 5C2. Specializations: attitudes and motivation in second language learning, ethnic stereotypes, statistics and research design.

Applied Language Learning
1991, Vol. 2, No. 1, pp.29- 68

Inside the Classroom: Learning Processes and Teaching Procedures

Leo van Lier
Monterey Institute of International Studies

Introduction

This paper proposes a theory of L2 learning that differs from accepted modes of theorizing in our field.[1] Differences fall in two areas:

1. The theory is based on data systematically gathered where most L2 learning is conducted: the L2 classroom.

2. The theory explicitly strives to take into account theoretical and practical knowledge from fields that are relevant to classroom language learning: linguistics, education, psychology, social psychology, anthropology, sociology, philosophy, and the theory of science.

If we can forget for a moment where our current theoretical knowledge base comes from, and consider the above points on their own merit, we will be tempted to conclude: "But of course! This seems a very appropriate starting point for an L2 learning theory." If we then, however, scrutinize our predominant theoretical frameworks, and list all the assumptions, hypotheses, and models we can think of, we will see that a very different picture emerges.

On the one hand, most current theoretical L2 work does not emphasize pedagogical[2] links, either as sources of data or as targets for application (see, Newmeyer & Weinberger [1988] for explicit proposals to separate theory and practice). This is not the place to refute such a separation of learning theory from actual circumstances of learning, but it does give those of us who are steeped in practice a warrant to "grow our own" theory, so to speak. This theory will be classroom-based and may eventually become a *theory of practice*, in the sense of Bourdieu (1977).

On the other hand, we notice that L2 learning theory charts its own empirical course pretty much without taking guidance from fields other than linguistics (particularly these days, Universal Grammar). A

glance at the index of virtually every L2 learning book or journal shows a dearth of interdisciplinary references and a failure to keep up with current trends in educational theory (itself a multiple-source field, and thus of broad use to L2 learning theory). This trend is changing somewhat, with psychological literature (Vygotsky's socio-cognitive view of language learning [1978], recent Piagetian work on consciousness [1975], Gardner's [1983] and Sternberg's [1988] new theories of intelligence, Deci's intrinsic motivation [1975; Deci & Ryan, 1985], and so on) and knowledge from other fields (note, e.g., the recent interest in C.S. Peirce's views on thinking, learning and scientific activity, 1955) increasingly being discussed at conferences and in professional journals.

In the new classroom-based theory that I will outline, the aims of the teacher and the researcher are no longer unrelated or opposed; they are identical, or at the very least convergent: understanding L2 learning by way of a close examination of actions and interactions that are undertaken with the goal of L2 learning in mind. Understanding learning is the essence of being a good teacher. And it is also the essence of a theory that is not based on a search for causal proofs (a search that has largely been abandoned by the social sciences; for a coherent discussion of "postpositivism," see Fiske and Shweder (1986).

The success of an L2 learning theory of practice will in large measure depend on our ability to relate intrapersonal L2 learning to interpersonal L2 learning or, in other words, to integrate cognitive and social aspects of learning. In this respect, our closest guide will be Vygotskyan psychology, but there are many other connections that will be traced as well.

To begin with, we will explore the road that L2 travels, all the way from *exposure* to *proficiency*. Once that is done, we will look at interaction between teachers and learners, and among learners, to find *opportunities for learning* (Allwright, 1984). Last, we will discuss ways to facilitate learning, by way of curriculum planning, course materials, and other avenues.

There are, thus, three tiers to our theory of practice (the major section headings of the paper are given in parentheses):

1. Theory of Learning
   (The Road to Proficiency)

2. Theory of Classroom Interaction
   (Analyzing Classroom Talk)

3. Theory of Curriculum and Teaching
   (Planning Improvisation;
   Innovation and Teacher Development).

The Road to Proficiency

In an ESL grammar class, I once asked students to bring an "entry ticket" to class every lesson (this was a 3x5 card on which they had noted, without worrying about the spelling, something they had heard in their daily pursuits). This request was based on my persistent suspicion that the students did not actually hear anything outside class–or rather that they did not listen systematically. I hoped that the entry tickets would force them to open their ears and start being aware of the real language around them.

Many small examples of language were brought into class this way–often idiomatic expressions such as "big deal!". Now, what does it take for a student to hear (notice) "big deal!" and to end up learning "big deal!"? This job is clearly not accomplished at once. From *noticing* the student needs to progress to *understanding*, at various levels, including the semantic and pragmatic. Eventually the student must learn to use the term in appropriate contexts.

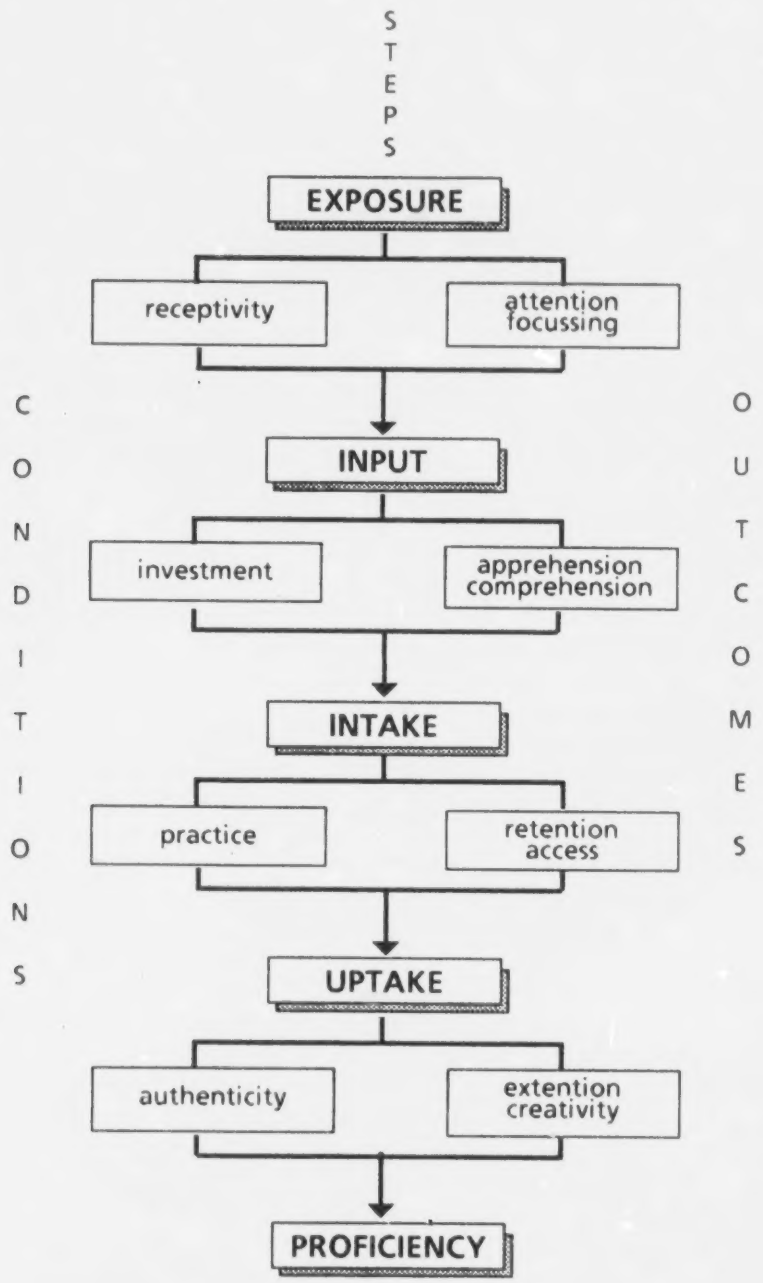
How is the complex task of learning, which involves many steps, and many different actions and skills, accomplished? For an answer, let us consider the road language travels from exposure to proficiency (by which I mean efficient use), as sketched in Figure 1.

Clearly, exposure is necessary to learn language. Less clear is what the ideal characteristics of this exposure-language are. Is quantity more important than quality? Are there limitations on the usability of exposure-language, e.g., that it should be comprehended, that the learner is receptive to it, that the learner makes an effort to process it, etc.? It is possible to imagine circumstances and environments in which any amount of exposure might be useless, and, conversely, situations in which small amounts of exposure might be sufficient. What, then, are the conditions for exposure-language to be usable for language learning?

Krashen and his followers stress that the language should be *comprehensible* (e.g., Krashen, 1985). People interested in Vygotskyan psychology focus on language that is in the *zone of proximal development*.[3] Yet others stress that meaningful interaction is the key that makes language accessible for learning. These different outlooks are not incompatible. Moreover by looking closely at the kinds of language that learners are exposed to, and the ways in which learners actively engage themselves with that language, we will undoubtedly find that one single, simple answer will not suffice.

One mistake often made is that researchers appear to assume that by looking at the exposure-language itself one can find the characteristics that make it either useful or useless. This is, of course, totally false. It is the learner, with her[4] prior knowledge, attitudes,

Figure 1
Learning: The Road to Proficiency



interests, analytical abilities, communicative dexterity, and so on, who determines the usefulness of the exposure-language.

Next, such issues as comprehensibility appear to be regarded as all-or-nothing features of language. Again, there is no reason this should be so. We comprehend things partially, just a little, more or less, a lot, but probably–at least in the case of language–never one hundred percent, and never instantaneously. Just how much we need to comprehend in order to be able to profit from the exposure-language is an open question, but it surely does not have an all-or-nothing answer.

Exposure-language is the language around the learner. It is available if the learner can use it and chooses to do so. As an example, an immigrant may work long hours in a small textile factory in New Jersey, every day traveling an hour-and-a-half from Astoria, Queens, to work. Along the way, and at various times of the day, she is exposed to a considerable amount of English, but she does not attend to any of it. Minimal instructions, greetings, and stock phrases carry her through the working day. Subway English is ignored. At home all interaction is in Spanish; for business English she relies on her savvier sister and brother-in-law, who live in the same house. One can say there is considerable exposure to English, but this exposure is, for whatever reasons, not usable.

When a person is *receptive* to the language she is exposed to, she will *attend* to it. This attention turns exposure into *input*, language that is received through the senses and made available for mental processing.

The learner then needs to *invest* effort (cognitive, emotional, physical), so that the input can be processed. Interactive work (conversation, for example) can assist greatly in this processing. In fact, we might call conversation "social processing" to match the term "cognitive processing." Investment, both social and cognitive, leads to two kinds of understanding: *apprehension*, a kind of intuitive, unanalyzed knowing (see Kolb, 1984);[5] and *comprehension*, including the more familiar types of conscious knowledge. These two types of understanding in learning go hand in hand, and both are indispensable. At this point, preexisting schemata can be brought into play, and schematic structures can be assembled to make sense of the incoming information. In addition, the new information is situated in the world and related to people and things. When this is done we can say that the input has been transformed into *intake*.

At this point, the language learner's job is not yet finished. The understood (apprehended/comprehended) material must be brought under conscious control. It must be accessible at the right time in the right way. This processing requires memory work, motor control, schematic networking, pragmatic matching (involving norms of appropriativeness and rules of use), and more. These processes occur through practice, a problematic notion we will return to later. The

result is successful retention. The term *uptake*, used in the diagram, reflects the aim of practice: to ensure that what goes in, stays in.[6]

The final stage is the free and creative use of the new material, under *authentic* conditions, when the new language is available as one option in the entire system of choices that the growing L2 affords the learner.

For the sake of simplicity I have shown this process in a linear form. In reality, the process is cyclical and spiral, with conditions and outcomes continually interacting with and reinforcing one another in a multidimensional sociocognitive space. What I have shown is the depth of the process of language learning–a step in the process of understanding the learner, probably the most crucial task facing conscientious teacher.

Focusing

Under the notion of *focusing* I want to subsume concepts such as "focussed attention" and "noticing-the-gap" (Schmidt & Frota, 1986; Schmidt, 1990), and, in general, any effort made by the learner to attend to particular features of language for the purpose of learning. This includes the familiar notion of focus on form. But I want to stress very strongly that I consider focus on language a much broader term than merely focus on form, which is just one aspect of it. I do not want to use the verb "focus" in a more transitive sense, as in "teachers must focus the students on X." This sounds too much like focusing a camera or binoculars, and I do not consider such mechanistic metaphors helpful. On the other hand, assisting students in their focusing efforts (including drawing attention to X) is, of course, a key task of the teacher. The success of such efforts depends in large measure upon the degree to which the teacher understands what the students are doing.

Focusing, then (see also Peters, 1983), is paying attention to aspects of language one is ready to learn, or at least is ready to have a shot at learning. In the past, teachers, textbooks, syllabuses, and test writers have decided for the students what was supposed to be focused on–in other words, focusing was externally imposed. The endless succession of methods and fads over the years indicate that none of those attempts are, in the long run, satisfactory. This has led to a current groundswell of opinion (Clarke, 1983; Danesi, 1988; Prabhu, 1990; van Lier forthcoming) that method is dead (let's hope that the next line is not "Long live method!").

Let us start from the reasonable assumption that the only efficient focusing is internally driven, by the learner's L2 system-to-date (known as Interlanguage) and by the continuing desire to learn and to engage in L2 interaction.

In L1 acquisition, I suggest, focusing on language is largely genetically driven. This focusing is often intense, all-absorbing and accompanied by a total readiness of the child to act upon whatever language material comes his way. It is similar to the heightened states of awareness called "vigilance" by Price in his classic philosophical discussion of perception (1969), as well as by Piaget in his later experiments regarding the role of conscious control in problem solving (1975). Price uses the example of a cat eyeing a nearby mouse and being ready to pounce at any moment, in any direction, whichever way the mouse might happen to move.

In L2 learning, this biological vigilance cannot be assumed, so attention must be fostered by other means–including making things interesting and comprehensible, setting high expectations, motivating the student, and all the other age-old pedagogical recommendations. *Language awareness* (Hawkins, 1987; van Lier forthcoming) and learning awareness (e.g., of strategies and styles; see O'Malley & Chamot, 1990; Skehan, 1989) or, in more global terms, language-learning awareness, are important elements in all kinds of learning, and a classroom-based theory of practice will need to pay more attention to them than have teaching methods to date.[7]

In L2 instructional settings attention-focusing needs to be fostered. But this does not mean we will always be able to predict on the basis of existing preconditions. what will be chosen as a target for focusing. Nor do we have to account for the myriad of linguistics items that are not focused on. To use Neisser's analogy: "To pick one apple from a tree you need not filter out all the others; you just don't pick them." (1976, pp.84-85). However, to continue the analogy, you are not likely to pick any apples that are out of reach.

The lesson for the exposure-provider is, it seems, that one need not worry about the exact nature of the language. Rather, one should provide variety and interest. Make sure there are lots of apples on the tree, and check that the learner has access to at least some of them. The learner will take what he needs and leave the rest.

Analyzing Classroom Talk

There are endless ways of using language–some public, some private. The importance of private speech is often overlooked, but it nevertheless plays a crucial role, both inside and outside the classroom (see Figure 2 , especially the activities mentioned under quadrant II, p. 38).

In the classroom, however, a prime way of using language is in speaking and listening to one other. This language interaction and participation in activities designed to enhance language learning

constitute the main reason to have language lessons at all. When we investigate the classroom language learning process, the characteristics of interaction and participation are therefore our main concern (even if we feel that private speech and other covert mental activities may well constitute the "real work" of learning).

We look at classrooms in order to *understand* the pedagogical quality of the activities that go on in them. I say understand so as to avoid saying *measure*. Measuring tends to mean, in the orthodox view of scientific method, the counting of units, and I agree with C.S. Peirce (1955):

> Now characters [Peirce means: characteristics, or qualities] are not units, nor do they consist of units, nor can they be counted, in such a sense that one count is right and every other wrong. Characters have to be estimated according to their significance. (p.152)

Measurement, in the sense of finding units to quantify, is therefore an inappropriate way of going about understanding the quality of interaction. What are the alternatives? We must examine classroom activities in terms of the learning opportunities they create for the students. In order to do this we need a learning theory. I sketched such a learning theory at the beginning of this paper in my description of the "road" from exposure to proficiency. Of particular interest are the conditions mentioned in Figure 1 (the left side of the diagram): *receptivity, investment, practice, and authenticity*. I will first discuss the twin concepts of receptivity and investment, which, together, make up much of what we generally mean by motivation.

Motivation: Receptivity and Investment

Within the view of learning expressed in Figure 1, the concepts of receptivity and investment are the most important prerequisites for quality pedagogical discourse. The former because it focuses attention and facilitates efficient perception; the latter because it promotes depth of processing and cognitive effort and efficiency. At the cognitive level, depth and effort are required so that new linguistic material can beprocessed and connected with existing knowledge and skill structures. The stronger and more varied the connections, the more likely new information will be retained and accessible. Receptivity and investment–and hence cognitive depth and efficiency–are most likely to thrive when the following conditions obtain in the classroom:

- The language is meaningful and understandable.
- The student is intrinsically motivated.

These two conditions require some elaboration. The connection between comprehension and retention (memory) has been well established for a very long time. Spinoza stated a few centuries ago: "The more intelligible a thing is, the more easily it is retained in the memory, and, contrariwise, the less intelligible it is, the more easily we forget it." (Spinoza, 1911, p. 254). And here is a more recent quote: "Children retain in their memory only what they have understood." (Piaget, 1975, p. 205).

Although we must bear in mind the caveat proposed at the beginning of the paper, that comprehension tends to be a "more or less" rather than an "all or nothing" affair, it is clear that meaning and understanding are key features of high quality pedagogical discourse.

Studies show motivation to be central to virtually all kinds of learning. In L2 literature, discussion has centered on two types: *instrumental* and *integrative*, with the latter in most circumstances being regarded as the most profitable (Gardner, 1985; this volume). A different kind of motivation, *intrinsic* motivation, has received much less attention in L2 research (see Crookes & Schmidt, 1989). Yet it is arguably the most important kind (Deci, 1975). Intrinsic motivation is the desire and commitment to excel in an activity because of the fun of it, the intellectual stimulation, or whatever–in short, the "worth-doingness" of the activity in itself. If the task is done for a reward or a threat at the end of it, then the motivation becomes *extrinsic*. Clearly most of what is subsumed under integrative and instrumental motivation is extrinsic in this respect. Aspects of such extrinsic motivation, especially if accompanied by strong societal expectations, a firm sense of commitment, and an environment of respect, can contribute significantly to intrinsic motivation. However, if the learning activities are dull, meaningless, distasteful, or otherwise perceived as not worth doing, it is unlikely that extrinsic factors, however strong and positive in general, will provide the impetus that is necessary to reach the required levels of receptivity and investment.

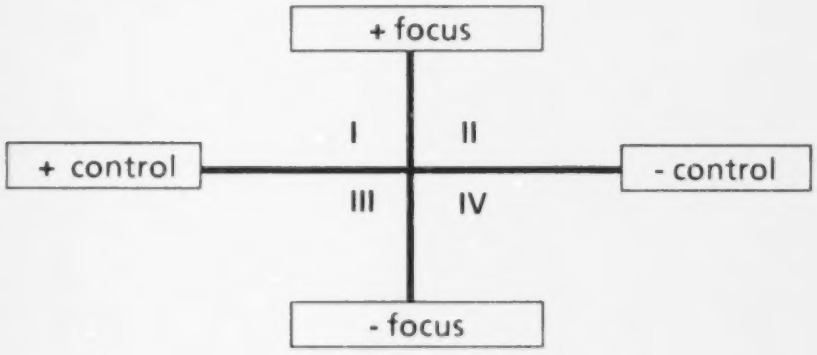
Readers who have seen the movie "Stand and Deliver," about the famous teacher Jaime Escalante, will be familiar with the term "ganas." It is a Spanish word that Escalante used to inspire his students, a word one also hears at soccer matches in Latin America. Whenever the players appear a bit sluggish the crowd shouts "Vamos muchachos, con ganas!" The closest equivalent in American English might be: "Come on guys, let's hustle!" Implying enthusiasm, energy, and commitment, the term ganas is close to what we mean by intrinsic motivation. Without it, it will be practically impossible to create profitable learning opportunities.

Language Practice: the Avoidance of Avoidance

Does practice make perfect? In the view of the language-learning process outlined here, the learner must be actively engaged with the target language in a variety of ways in order to ensure sufficient cognitive effort and depth of processing. But what kinds of active engagement are beneficial? According to Figure 1 practice is required to move from intake to uptake. But what kinds of activities do we call practice?

Ellis (1988) takes a detailed look at what we mean by practice. He concludes that, given the variety of definitions, it is no surprise that quantitative studies have produced very mixed results. He proposes that practice can vary along a continuum from (externally) *controlled* to *free* ("more control" means that both language and actions are more predictable), and that it can be more or less narrowly focused on specific linguistic features. As an initial conceptualizing exercise it is useful to take these two continua of *control* and *focus*, and construct a simple semantic wheel, as follows:[8]

FIGURE 2
Types of Language Practice



Below is a tentative description of the resulting four types of practice, with some likely exponents:

I. controlled and narrowly focused practice; this category will include many varieties of mechanical drills, grammar (fill-in, and so on) exercises, translating sentences, etc.

II. focused but not controlled; I would probably take a broader view of focus here than Ellis does, in line with our discussion in the preceding section, and include under this heading the very important work of inner speech, language play, private rehearsal, planning, and so on. Much of this work may be covert, and therefore easily escapes notice in empirical studies.

III. controlled but not (narrowly) focused; this category might include such types of practice as guided dialogues, role taking, simulations, certain information-gap tasks, and so on. The less controlled such activities are, in the sense given above, the closer they move towards IV.

IV. not controlled and not focused; this category means communicating per se (i.e., conversation), and might not normally be regarded as practice, although of course it is a way of practicing.

As regards controlled practice, we can note Ellis's conclusion:

> We are led to conclude that in the case of controlled practice the old axiom "practice makes perfect" may not apply to language learning or, at least, not in the way that many teachers and methodologists think it does. Practice may only facilitate acquisition directly if it is communicative, i.e., meaning-focused in nature. (1988, p. 38)

If careful study were to confirm this conclusion, we might be tempted to deduce that practice types II and IV are the only beneficial ones. However, a few strong qualifications are in order. First, there appears to be no reason to assume that controlled work can never be meaning-focused. Second, some kinds of controlled practice might facilitate L2 learning indirectly, in important ways that we may not have identified to date. Third, both control and focus are continua, and there may be optimum amounts of both, so that too little control or focus may well be just as detrimental as too much. For these reasons it is somewhat premature to draw conclusions.

There are two issues that are directly affected by the amount of control and focus of language practice activities. Less control and less focus allow for more and freer interaction, or negotiation of meaning, as it is often called, whereas more control and more focus are more likely to put the spotlight on error correction or, in more general terms, *repair*.

Negotiation of meaning creates useful learning opportunities since it makes activities more meaningful and enhances understanding. Furthermore, it provides opportunities for listening and speaking in context. For negotiation to be possible there must be some information that is not given to all participants, but that has to be found and/or

shared. In the literature, this feature of tasks is called an *information gap*. Extending this notion, Prabhu (1987) makes a useful distinction between information gap, reasoning gap, and opinion gap activities. However, we may go further yet and note that all such gaps refer to a *transactional* use of language, since they imply a transmission of information and ideas. In addition, language is used for *interactional* purposes, that is, for the creation and maintenance of social relationships, the presentation of self, and other social concerns (see Brown & Yule, 1983). In our constant search for gaps we may easily overlook this crucial social function of language. Later on in this paper I will illustrate some ways in which this function is incorporated in lessons.

When the language and interaction are to a significant extent controlled, and there is a focus on language, correction of linguistic errors is likely to occur. Elsewhere (van Lier, 1988a) I have argued that error correction is a type of *repair*, which is a more general activity of solving problems in interaction, and which includes a number of things which might be subsumed under negotiation of meaning, such as requesting clarification, checking for comprehension, and so on. Such repairing can be done by the speaker (self-repair), by the hearer (other-repair), or by both, in cooperation (collaborative repair). One way of doing such repairing is by correcting errors. However, correcting errors, if done by a listener, tells the speaker that he did something wrong. We all know that if we do this to friends, relatives, or colleagues, we tend to upset them, to cause trouble. We therefore avoid such other-repairing, and work towards self-repair as much as possible (Schegloff, Jefferson & Sacks [1977] call this "the preference for self-repair"). In classrooms things might not be so delicate, but we must nevertheless not assume that normal social concerns are suspended altogether. Further, we must remember that a proficient language user carefully and efficiently monitors her own language and context of speaking in order to do self-repair when required. In any ordinary speech event you are on your own in this respect. Students must therefore not be made dependent on teachers by constant other-correction, since this would mean that they would not learn to self-monitor efficiently and to repair their own problems.

I do not wish to imply that error correction is inappropriate at all times in the classroom. Rather, we should see the issue of *type of repair* as a function of *type of activity*. That is, what we are doing, and the reasons for doing it, will determine how we handle the issue of repairing problems. For example, if at some point we decide to have students produce correct pieces of language, then we must correct incorrect pieces of language during such an activity. If this correction produces undesirable results, then either the activity is unsuitable, or the correcting was badly done. From this perspective, it is unhelpful to look at the

issue of correcting in isolation from the activities in which the correcting takes place.

I want to conclude this section on a more general note regarding the utility of L2 classroom work, whether we call it language practice or something else. All this work has one straightforward function, which we sometimes tend to forget in our communicative exuberance. Whereas in daily life all sorts of demands, concerns, worries, excitements, and so on, may conspire to make us play it safe and avoid uncertain aspects of L2 use, this avoidance is unnecessary in the classroom. There the learner enjoys the luxury of trying out "unsafe" things, goofing around with unfamiliar linguistic items and expressions, and in general pushing the frontiers of interlanguage forward.[9] Thus, in a very important sense, language practice in the classroom allows the student to "avoid avoidance."

Authenticity

S1 What's mean leave-taking?
T Leave-taking. You say "goodbye". OK? Sometimes when you leave somebody you say "goodbye" or "see you tomorrow." Or what else can you say?
S2 See you later.
T See you later, yes. What else? Huh?
S3 Take it easy.
T Take it easy. What else? S4?
S4 Catch you later.

The above piece of language[10] is quite easily identifiable as L2 classroom language. Even people who are not language teachers or students would probably be able to correctly place it at once. Why is that? And what does such an observation mean?

For many people the fact that classroom language looks and sounds like classroom language is evidence of the artificiality of language lessons. The inevitable recommendation that follows is that teachers must strive to make their lessons more "natural," since otherwise students do not get realistic and meaningful practice. For example, at a recent conference I heard statements such as "ESL classroom language is about the most artificial language in American society; I firmly believe that." (quoted verbatim, though I see no point in identifying the locutor). However reasonable such arguments may sound, I think we would do well to examine them carefully and critically. Please read the following exchange which, I admit, I have just made up:

A: Open your mouth wide, please. Hold still for a

moment, if you can.
B: ((mumbling through multiple gadgets)) Ish ick gonna hur"?

Where does this take place? At the dentist, of course. Few people would mistake this for an interaction at a fast food restaurant. Yet, no one complains about the artificiality of the language (though there may well be other things to complain about). The point is that every speech event in a setting shows features that tie the speaking to the setting, usually in such obvious ways that it is immediately clear (when reported) where the speaking was situated.

There may be many things wrong with classroom language, but the fact that it is identifiable as classroom language is not one of them. That is an obvious and uninteresting observation in itself. But we have to investigate classroom language with more important questions in mind. How does it indicate learning processes at work? How does it facilitate those learning processes? These questions are related yet quite distinct for the researcher (and I include the teacher who, after all, must be a researcher in order to understand her students). And they state in a nutshell the complexity of the work that needs to be done if we want to understand and improve L2 learning in classrooms.

Simplification, Foreigner Talk, and Teacherese

A central notion in any form of communication is *audience design* (Bell, 1984; Levelt, 1989). In general terms, whenever we communicate with another person, we are concerned that he understand what we have to say (exceptional purposes aside, such as deliberate obfuscation to befuddle or impress the interlocutor). For example, as a linguist I may be able to lecture at length on topics such as tone sandhi, defective verbs in Russian, ergativity, and verb measures in Arabic, but I am unlikely to do so at a neighborhood party, even if the conversation were to turn to language.

When we talk to children, we adapt our language, and when we talk to foreigners we do likewise. The more we understand the problems our less-competent interlocutors face, the more efficient we are likely to be at adapting our language, and at deciphering what is said to us in return.

Note that, in changing our language for the benefit of a foreigner, we are practicing some form of *accommodation* (more precisely, convergence; see Beebe, 1988, for more detailed discussion of these terms). If we do not like the foreigner, or harbor some form of resentment for whatever reason, we may deliberately block this accommodation, with the result that the foreigner does not understand what we say, and vice versa (divergence). Stereotype has it that

Parisians do this to tourists who speak French less than perfectly. There may be other reasons why foreigner talk could be unsuccessful: Perhaps we have insufficient experience with non-native speakers to know *how* to accommodate. The usual caricature is that of a country bumpkin (well, perhaps urban chauvinists should be included as well) merely repeating what he said before, only louder.

An L2 teacher is supposed to know the kinds of problems non-native speakers have. Furthermore, she is mandated to understand the particular students in her class, and can therefore be regarded as especially sensitive to and aware of what they know and can cope with. In the classroom this leads to a paradoxical situation where, on the one hand, the teacher is uniquely able to adapt her language to accommodate the students' imperfections but, on the other hand, by doing so deprives her students of the full array of language and expression that native speakers engage in. The more the students depend on the teacher in terms of exposure and modelling of language, the more acute this paradox will be. After all, in many settings students will be able to hear much L2 outside the classroom, but in many others they may not, and this must be reflected in the classroom.[11]

The characteristic features of teacher talk have been described in many studies (see Ellis, 1985; Chaudron, 1988). Here, I want to point out a danger that I have noted with many methodologists' interpretations of *comprehensible input*, a notion familiar from the work of Krashen. It is particularly prominent in ESL and bilingual education workshops (though there may well be parallels in other settings and for other L2's). I take my criticism from numerous handouts and training manuals, though it would serve no purpose to cite them and trace them to their origins. In order to provide comprehensible input, the advice goes, the teacher must use exaggerated intonation, avoid using pronouns, speak slowly, and so on. Even though it may very well be the case that teacherese generally has the characteristics described, I would suggest strongly that the teacher beware of recipes for using simplified and distorted language. As a general policy such prescribed standardizations of teacher speech are unnecessary, counterproductive, and probably insulting to most audiences.

Rather, I would like us to ask ourselves: Now that we are conscious of our tendencies to engage in teacherese in the classroom, how are we going to control it? By "controlling" it I do not mean eliminating it–but rather, bringing it under the control of conscious pedagogical decisions, so that we can use teacherese to its best effect, without overdoing it or neglecting it.

The questions that are relevant in such a self-monitoring exercise are these:

a. Do I use teacherese, and what is it like?

b. Should I use teacherese more or less? Why?
c. Should the quality of my teacherese be improved? How?
d. When is using teacherese appropriate, and when is it not?

Although I do not wish to deny importance to questions (a) through (c), in my view question d) is the most interesting one here. It suggests that teacherese can be brought under the conscious control of the teacher, and become a *strategy* to obtain certain outcomes. Most importantly, it suggests that the teacher can use more than one register, or linguistic repertoire, in the classroom, thus increasing the range of expression available as exposure to the student.

In sum, rather than being told that teacherese is necessary in order to provide comprehensible input, and then being given instructions as to how to do it, we will investigate our own teacherese and learn to use it as one tool among many. It thus becomes part of the theory of practice we are creating.

Types of Authenticity

When we use teacherese we bring into the classroom a type of language specially designed to make the job of learning easier for the students. But it is language that is different in some respects from the language as it is spoken in most other places. Moreover, it is not only the teacher's speech that is changed in order to facilitate learning: Books, tapes, worksheets, and other materials may contain language specially designed for the benefit of the students. I argued above that in itself there is nothing wrong with this. However, if the audience-design results in language that is artificial and of a narrow range, one might reasonably object that students do not get realistic exposure and practice. The question is, how can we satisfy the natural demands for audience design, while at the same time providing a rich and authentic linguistic environment?

The notion of *authenticity* has been prominent in L2 methodology for a long time (see Widdowson, 1979; Breen, 1985; Cathcart, 1989). Yet I feel that the issues have rarely been more than partially discussed. In essence, we must never lose sight of the crucial task of finding a balance between two legitimate, but apparently contradictory, concerns: facilitating understanding and learning on the one hand, and providing a rich and undistorted linguistic environment on the other. An authentic L2 classroom is one that achieves this balance, not one that only addresses the latter concern.

Traditionally, authenticity has been taken to refer to the language materials used in lessons. For example, many textbooks used to have (and still do, of course) dialogues that were rather odd, because all their authors wanted to do was to highlight certain linguistic features. A textbook of Turkish I came across recently had such memorable exchanges as:

A: Who has three eyes?
B: No one has three eyes. Everyone only has two eyes.

Few teachers would disagree that some improvement in such language might be beneficial. However, imagine that the textbook writer responds, in the new edition, by changing "eyes" to "books" or "'dollars"? Would that be more authentic? If so, why? If it still sounds weird, why?

What I am getting at is that there are infinite degrees of authenticity, as well as several types of authenticity. I cannot describe all of these in sufficient detail (see van Lier, forthcoming), but I will just briefly indicate what they are.[12]

First there is authenticity in the sense in which it is used when we talk about antiques or works of art. When we say "Is this painting authentic?" we mean " Was it really painted by X?," or "Is it really what it claims (it is claimed) it is?" This we can call *authenticity of origin* (or production, creation).

Second, there is a sense in which language is authentic if it "fits" the context in which it takes place (e.g., in the way that the dentist example above fits the context of a dental appointment). This is *authenticity of context*. It includes genuineness of social interaction, and effective audience design.

The third type of authenticity relates to the purpose of the user, and whether the language is suited to that purpose (even if this includes a hidden purpose). This is *authenticity of purpose*. It includes, as a special subtype, authenticity of representation or reproduction, when the presenter of the language (e.g., the teacher) is not the creator of the language (e.g., a newspaper reporter, a poet). To illustrate, the latter is a perennial issue in the performance of classical music on modern instruments and directed by conductors who may interpret the composer's score in their own way.

Together, authenticity of context and purpose can be called *pragmatic authenticity*. Language that fails the various criteria falling under the heading of pragmatic authenticity is distorted (cf. the use of this term in Habermas, 1987) or defective. I would call excessive simplification or teacherese as discussed above defective or distorted in this sense. A distinction must be maintained between appropriate

accommodation and inappropriate distortion. That this demand is difficult to define and balance does not make it any less important.

The fourth type of authenticity follows directly from the two aspects of pragmatic authenticity. It deals with the pedagogic context, requiring that language used in a pedagogic setting should be *pedagogically authentic*–that is, it should not deny that it is about teaching and learning (note that a "let's pretend" kind of language use can be pedagogically authentic within this requirement; but if we act and speak consistently as if this lesson were not a lesson, and I am not a teacher, and you are not students, then I would say that we violate a principle of pedagogic authenticity).

Fifth, and finally, there is the question of intellectual interest in, moral commitment to, and emotional involvement in the work at hand. This is arguably the most important criterion, one without which all the other types of authenticity are useless (it is closely linked to intrinsic motivation, as discussed above, as well as to autonomy [Deci & Ryan, 1985]). For want of a better term, I call this *existential authenticity* (see Magee, 1973, p. 68, for a similar use relating to scientific activity).

If we see the issue of authenticity in this light, we do not have to limit our judgments to the usual question of whether or not the materials were produced for purposes other than language learning, or if the activity is one that resembles activities that go on in the so-called "outside world." Rather, our question now is:

> Am I using undistorted language sincerely to further language learning, to promote commitment and interest, and in ways that make my intentions clear to my audience?

This means that language specially made up for the occasion, and activities invented for the classroom, can be authentic in far more valid and meaningful ways than some article clipped from the newspaper, a videotaped conversation, or a student-student exchange about "what I did last weekend."

Planning and Improvisation

Teacher–training courses spend considerable time on the planning of lessons and programs. Supervisors usually want to see the teacher's lesson plan, and also hope to see some sort of relationship between the lesson they observed and the plan for that lesson. Planning a lesson (and carrying out the plan effectively) is therefore a valued part of a teacher's professionalism.

We also know however, that things don't always turn out as planned; when the unexpected does happen, we must be prepared to improvise. This improvisation is also part of a teacher's professionalism.

Through classroom research we can investigate the quality of interaction (e.g., the occurrence of learning opportunities) during planned and improvised parts of lessons. We also can find out what proportion of a particular lesson is planned, and how much is improvised. Next, we can see if recurring problems happen primarily in planned or in improvised portions. As an example, the extract on p.41 is an improvised classroom episode: A student asks an unexpected question. In response, the teacher gives a brief paraphrase, a couple of examples, then asks several students to give more examples. I assume we can call this a successful bit of improvisation.

However, this improvised episode may have occurred inside a larger planned activity, e.g., a textbook exercise on greetings, goodbyes, and other conversational formulas. This suggests that, in any lesson, planned and improvised actions and interactions may be tightly interwoven, and that a teacher may have to skip back and forth between the two. Sticking too rigidly to the plan may be just as detrimental as having no plan at all–especially if it turns out (as I suspect, and will propose later) that the most profitable learning opportunities occur in improvised portions of the lesson.

In this section I want to discuss two classroom issues that relate directly to planning: *scaffolding* and *learning materials*.

Scaffolding

The pedagogical notion of scaffolding is borrowed from the work of Jerome Bruner, who investigated the interaction between parents and children. Bruner focused on the "peekaboo" game; see, e.g., Bruner, 1983; Cazden, 1983; van Lier, 1988a; for early suggestions to apply the scaffolding metaphor to classroom pedagogy.) "Scaffolding" means setting up a situation in such a way that participation is easy and mistakes can be rectified (the *scaffold*), and then gradually handing over parts of the role, the task, and the interaction, as the students show they can handle them (the *handover principle*).

In order to illustrate pedagogical scaffolding in L2 classrooms I include a transcript of an ESL grammar lesson.

> The setting is an ESL grammar class with about 10 students, all Asian. "T" wants to use the overhead projector (which is on a small table) and is enlisting the help of student S2, as well as asking other students to give directions (practicing phrasal verbs).

1T . . . S2 . . can you help me? . . . . . who wants to tell S2 what to do? remember the verbs we practised last time? you want to all help him? . . and give hime some instructions? . . . . .
2S1 ((softly)) (plak it in) . . .
3T say it again?
4S1 plug it in
5T (relu-) . ahah plug it in. Okay, we'll plug it in. . . . . plug it in. and now?.
5S3 Turn it on
7T turn it on . . ((to S2:)) you just do what you're told okay, just do what they tell you . . . nononono . . ((to S3:)) what did you tell him?
8S3 turn it on . . . . . turn it on
9 ((click))
10 T turn it on . okay. next?
11 S5 around
12 S3 ⌈turn around
13 SS ⌈turn around//around//((unint))
14 T pardon?
15 S5 turn around
16 T turn- . .
17 S5 ⌈around
18 T ⌈it . . . around
19 SS ⌈((unint))
20 T turn it around
21 SS ((unint))
22 ((noise of OHP being turned around))
23 T o::hkay . . ya . . has he finished? is that every what else does he have to do?
24 S1 focus . . .
25 S ⌈focus
26 T pardon?
27 S1 focus
28 T focus. okay. focus . . . . . .

This activity shows scaffolding at three levels, going from "macro" to "micro," and from planned to improvised. First, we can note that the activity of setting up the over-head projector (OHP) is used to get the students to practice phrasal verbs in English. It is important to realize that this is not a one-off task, but rather the fourth or fifth such task in an indefinite series spread out over the entire course. This series is scaffolded in the same way that the peekaboo game is scaffolded over several years of a small child's life: First the parent does all the "peeking" and "booing" (or whatever we wish to call the actions in the

game), the popping up and disappearing. Gradually the child takes a more active part, at last becoming able to initiate and sustain the game herself. In the same way, the OHP activity is at first entirely conducted by the teacher, but gradually the students play a more active role, giving instructions and performing the actions. The illustrated episode may be about midway through the cycle: The teacher still directs and prompts, but the students already give instructions and carry them out. The OHP activity can thus be projected over an entire course of studies, and can be varied by using whatever other technological aid is actually used in the lesson: a VCR, a computer, a cassette recorder, and so on. In the lesson plan, however, it may be merely indicated briefly as: "OHP version 5; approx. 10 mins."

The OHP task is also scaffolded at a more detailed level. We note that it consists of several actions that must be performed in sequence (e.g., the projector must be plugged in before it is turned on). This sequence has been worked out in advance by the teacher, partly because he wants to get as much mileage from the task (in terms of different phrasal verbs) as is feasible. Turn 7T shows this scaffolding at work when the "helper" wants to do several things at once without having been specifically instructed by his classmates. The teacher intervenes at that point to re-establish the sequence.

The third way in which the OHP task is scaffolded occurs at the interactional level. Here the teacher decides when to prompt, help, and correct the students to get the job done, trying to ensure that it is neither too easy nor too difficult for the students.

Scaffolding, as illustrated here, is a three-layered teaching strategy consisting of actions, and sequences of actions that are partly planned and partly improvised. At all three levels it focuses on understanding what is easy and what is difficult for the student. It allows the teacher to keep in mind, at any moment in the lesson, both a long-term sense of direction and a short-range plan of action. This *dual vision* (see Dewey [1904] for a related notion) helps in making the right split-second decisions that are necessary at every step in the lesson, however well-planned it may be. Thus, pedagogical scaffolding creates a balance between planning and improvisation.

The Role of Materials

In L2 classroom learning it might well be said that "the medium is the message" (after Marshall McLuhan's teachings, popular in the '60s). After all, if the L2 is used as the medium of instruction (as it should, as much as possible), the means and the goal of instruction are one and the same. In that sense, anything that is talked about in L2, including, e.g., the OHP in the task illustrated above, is "teaching material." A conversation, an impromptu joke, an announcement about

a picnic, a set of instructions, and so on, are all teaching material that can be "exploited" without actually bringing anything into the room. Since this classroom language is "just there," we tend to forget about it and perhaps not use it efficiently (e.g., unnecessarily use L1), and this is a pity since it is probably the most authentic language (in a pragmatic sense, at least) available in the classroom. Moreover, given its context-embeddedness, it is likely to be understandable and relevant as it is, without modification (see further below; see also van Lier, 1988b).

If we did nothing but talk about classroom things and whatever else came to mind, however, our lessons would probably be rather dull and uninspiring. There are two types of materials, in addition to here-and-now classroom realia, that can be brought into the room. First, information about the language; second, information about all sorts of other things (notice at once that these two types are not mutually exclusive or incompatible).

As far as language materials are concerned, we will be mindful of the never-ending debate of *focus on form* versus *focus on meaning*. It appears, from the methodological literature, that you focus either on one, or on the other. Further, it is assumed that if you focus on meaning you do not explicitly talk about language at all. I think that somehow we have got our categories mixed up here: Form and meaning are both part of language, and a focus on language embraces both form and meaning. We do not therefore have to choose between form and meaning; rather, our question becomes: Do we focus explicitly on language or not? If we say yes to this question, we can then discuss the balance between form and meaning that appears most appropriate. If we say no, we opt for an uncompromising content-based approach, though even here we cannot control how much focusing on language goes on inside the student's head. In general, it appears, we do not know for sure if explicit language work is essential, to what extent it can be beneficial, for what kinds of students and in what situations, and so on—though limited research suggests that a focus on language is beneficial (Oskarsson, 1973; Swain, 1988). Furthermore, Ellis's suggestion (1985) that explicit language work can have a powerful delayed effect makes intuitive sense to me as a language learner with a long and varied history.

L2 Textbooks

The bulk of materials used in L2 classes has been especially prepared for L2 teaching, and there is a long tradition of writing and publishing in this area, especially in English as an L2. From my own school days I remember the course books in English, German, and French as invariably being dull and unimaginative, and solidly grammar- and vocabulary-based. Although there were jokes, anecdotes,

and sometimes even drawings, it all reminded one of a stuffy uncle whose feeble attempts at entertaining the family had to be reluctantly endured.

It may be useful to classify L2 teaching materials as follows:

a. Full-scale courses, usually consisting of a series from beginner to advanced level, including student books and teacher guides, tapes, workbooks, and perhaps picture sets, videos, or other visuals. The philosophy of such series is to give the students all that they might conceivably need, to be self-contained, and to offer a clear progression of objectives and achievements over time. I remember seeing many such series in different countries, and it may be edifying to quote a particular series that had an interesting set of titles:

1. Here We Are
2. Here We Are Again
3. Here We Are Once More
4. We Are Still Here
5. We're Nearly There
6. We're There At Last!

One can almost hear the sighs of relief when book six finally comes in sight.

It is clear that such series need to be organized in some way, and much L2 methodology over the years has focused on principles of course organization. A particularly clear (and classic) representation can be found in Mackey's bible of the '60s *Language Teaching Analysis* (1965):

selection (what to choose, according to quantity,
proportion, and utility, in the areas of grammar,
phonetics, vocabulary, semantics);
gradation (grouping and sequencing of items selected);
presentation (demonstration, four skills);
practice (repetition, variety);
evaluation (testing, measurement).

Since the greening of our field in the '70s it may be that we have progressed from a more or less haphazard collection of linguistic items to a (less haphazard?) collection of functions or notions. But it seems to me that the basic organizing philosophy of the new courses that we see every year at conference displays has remained remarkably unchanged since Mackey's analysis. And there may be nothing wrong with that, of course, though some possible shortcomings will be pointed out below.

b. Supplementary books, which offer more specialized practice or guidance, perhaps in one of the skills (reading, listening, writing, conversation), or catering to a particular profession (French for engineers, Swahili for Peace Corps volunteers), or using particular techniques (pictures for problem solving, games, role plays). My all-time favorite title for an L2 supplementary book, a remedial book for English published somewhere in Southern Europe, is *Brush Off Your Broken English*. Such materials give a teacher more choice and flexibility, but in general they are used alongside a general text as padding, since in themselves they do not give sufficient backbone and direction.

c. Grammars, grammar practice books, and dictionaries. In general, this class of materials is meant to be used by students on their own, at home, for reference and individual practice, as needed. As Newmark said, as long ago as 1966:

> The student's craving for explicit formulization [sic] of generalizations can usually be met better by textbooks and grammars that he reads outside class than by discussion in class. (Allen & Campbell, 1972, p. 42)

We probably all know the student who comes into class armed with several dictionaries, grammars and other paraphernalia and who, unless restrained, will spend most of his time and energy looking things up. This has led to the banning of such books in many a classroom. Further, I can remember no more deadly practice (except for, perhaps, reading around the class) than doing grammar exercises consisting of long lists of unrelated sentences to which certain things must be done. For one reason or another, many teachers find that they have to include such exercises in their lessons. In my experience, such work only becomes palatable if frequent leavening is inserted–for example, by making personal comments or jokes on the basis of exercise items, as shown in the following extract:[13]

1 S2 'which' - a person or things
2 T yes .. yeah which dog is yours ... okay?
which baby is yours ... okay?
3 S2 I don't have baby
4 SS ((loud laughter))
5 T are you sure?
6 SS ((loud laughter))
7 S6 I don't kno-ow

If the grammar exercise is used as a springboard for conversation, jokes, puns, wisecracks, and so on, it may even become an interesting and

valued part of the lesson. We must realize, however, that whatever the students learn is probably learned in between the sentences of the exercise. "Learning between the lines," we might say.

d. The final category of L2 teaching materials contains probably the most interesting and promising development of recent years. I refer to the rapidly growing range of handbooks for teachers on many teaching-related topics (e.g., the teaching of reading, listening, grammar,pronunciation, and so on). I have found many of these books extremely useful and imaginative, and an armful of them is the most valuable baggage the teacher can carry around. If only we can somehow find the security and direction we and our students need without having to rely on an all-inclusive textbook series, this will be the format of the future. Even technologically advanced media, such as interactive video and computer programs, are nothing but textbooks and supplementary books on a screen–unless they offer some form of "open-stack" access which goes beyond whatever happens to be on the disk or tape. After all, it is very hard to see how a truly self-directed (autonomous) learner can develop if the materials continue to control the choices that are available.

The Curriculum

The perceptive reader will have noticed that the issue of L2 materials is closely related to that of planning. Teaching materials can assist in scaffolding at any one of the three levels described, but the calibration has to be made *in situ*, with teacher and students articulating what is difficult and what is easy. Planning and improvisation therefore have to be in balance, and it is hard to see how materials, especially textbook packages, can do this *a priori*. Educational endeavours, of whatever type, tend to want to predict what the outcome is going to be, and how it is supposed to be achieved. And the easiest way to predict this outcome is by way of carefully articulated curricula. In the view of learning and of pedagogy that I have sketched in this paper, such prediction is not feasible. Learning cannot be mandated, however "macho" (Krashen, 1990) the syllabus is. The same point is made by Prabhu, who prefers *simple* to *sophisticated* syllabuses. About the latter he says that:

> ...the process of enriching syllabus design can also be one of impoverishing classroom language and that the more 'content' a syllabus has in the sense of 'detail', the less exposure to language the learner is likely to get. (1987, pp. 92-93)

Stenhouse has said that the curriculum has to be brought into the room on "a porter's barrow" (Rudduck & Hopkins, 1985, p. 67), since it includes all the teacher knows and has experienced, and everything that influences his philosophy and practical actions. It forms the back-drop against which the business of classroom teaching is carried out.

The syllabus is a much more concrete, finite plan of action and a set of tools, including materials. My suggestions (and Krashen and Prabhu appear to concur) is that the curriculum should be vast and rich, but the syllabus simple and skeletal. This allows the teacher to pursue pedagogical scaffolding and to allow space between the lines for learning to occur.

Functions of Language in the Classroom

We have seen earlier that language practice can be characterized according to a greater or lesser focus on specific linguistic features. In our discussion of teaching materials we also noted that very often in L2 classrooms the medium is the message. In other words, if language is used to focus on language, our activities are *medium-oriented*. If, however, we are focusing on meaning, on exchanging information or on developing social relationships, then we might say that our activities are *message-oriented*.[14] A third possibility is that we are using language to organize things, to tell people what they are supposed to do or how they ought to behave, and so on. In that case our activities and language are *activity-oriented* (for more discussion on these orientations, see Ellis, 1984; van Lier, 1988a, 1988b).

These three orientations refer to the general sphere of interest we are addressing. In addition, we can use language to do different things, or fulfill several functions, such as (1) telling people facts or experiences, (2) working on relationships with other people, and (3) signalling connec-tions and boundaries, clarifying what we are saying, summarizing, highlighting, and other metalinguistic work. This gives us another tripartite division–this time Halliday's major language functions: the *ideational*, *interpersonal*, and *textual* functions.[15]

Each of these three functions (or macro-functions) can be used in each of the three event-orientations mentioned above. If we create a grid out of them, we can plot a number of different kinds of activities that can occur in an L2 classroom (as elsewhere, of course), and that can be examined for their potential to generate learning opportunities. This grid is presented in Figure 3.

The purpose of the diagram is to show the many kinds of activities that can take place in the L2 classroom. It can be used in planning or monitoring, to check if we are fully exploiting the potential for varied language use that the classroom offers. It is interesting to note that the controlled practice discussed earlier finds no place in this grid (except

perhaps as a sort of "outgrowth" of the ideational/medium box)! This absence certainly seems to be compatible with the things that Newmark and Ellis have told us (see above).

FIGURE 3
The Classroom Construction of Learning Opportunities

EVENT-ORIENTATION / FUNCTION	TOPIC		ACTIVITY
	MESSAGE	MEDIUM	STRUCTURING
IDEATIONAL	Information Problem-Solving Transaction	Explanation Rules	Instruction Pre-task
INTERPERSONAL	Social interaction Conversation	Role play Dialogue	Discipline Cooperation Songs Games
TEXTUAL	Meta-Cognition Critical thinking	Ethnography Meta-language	Learner training Meta-task

Innovation and Teacher Development

In this section we will begin tying together and exploring the significance of some of the major points raised in this paper. We will first look at the structure of lessons and the relationships between tasks and content. After that, I will propose some ways in which action research and self monitoring of lessons can improve teaching and make teacher development more relevant to the actual demands of L2 classrooms.

The Well-Formed Lesson

The term "well-formed" is usually used for grammatical sentences, but I use it here in the sense of well-formed discourse, following Stubbs (1983), who argues that, if discourse has structure, it can be well-formed or ill-formed. To determine the "well-formedness" of a lesson, we must, therefore, look at its internal structure–as well as place it in larger structures, alongside other lessons. Well-formedness is thus distinct from direct methodological judgments of quality (though

there are significant meeting points that will eventually become visible).

As far as the internal structure of the lesson is concerned, we have seen that it contains planned elements as well as space for improvisation. In that respect, it is similar to life in general, where predictable and unpredictable events occur in balance (ideally, anyway). Predictable events are routines and culturally instituted rituals that help us cope with the unpredictable that is bound to happen when we least expect it. In the classroom, too, we need such predictable activities as *points of stability* to help us handle unexpected and freewheeling episodes. I have suggested above that most advances in terms of learning are made during the improvised rather than the planned parts of a lesson. In addition, it is likely that most methodological advances in terms of *innovation* are also made during such improvised activities. In order to profit from such opportunities for learning and innovation, it is essential to monitor what goes on (see below). Moving from lesson-internal well-formedness to the more global view, every lesson should take into account earlier lessons, as well as prepare the ground for further lessons. To make the progression more sophisticated, recent approaches to syllabus design have replaced the old-fashioned linear format (one thing after another) with spiral and cyclical designs, in which material is systematically recycled and re-introduced in ever-more-complex contexts. But the progression that I have in mind–which weaves together points of stability based on a simple textbook, scaffolded activities, and improvised social interaction–is better expressed using the metaphor of Wittgenstein's rope:

> What ties the ship to the wharf is a rope, and the rope consists of fibres, but it does not get its strength from any fiber that runs through it from one end to the other, but from the fact that there is a vast number of fibers overlapping. (1958, p. 87)

It is important to note that in this process view of the syllabus it is not the textbook that "drives" the course, since the book is only there to provide points of stability in the lesson, and perhaps individual practice at home. Rather, we advance on several fronts simultaneously; we weave strands, and when one strand is played out, we weave in a new one.

For this way of working, a degree of flexibility and continuity are required that is not always compatible with the 45-minute lesson. We may be engaged in activities whose momentum goes beyond that short and arbitrary time span. Instead, we should think in terms of blocks and modules (in which themes, for example, can be explored).

Tasks and Content

Two of the most popular current innovations are *task-based* (Long, 1985; Candlin & Murphy, 1987) and *content-based* (Mohan, 1986; Brinton, Snow & Wesche, 1989) language teaching. Both orientations fit in well with the classroom learning theory I am outlining, and they are, of course, not mutually exclusive–in fact, one would probably fail without the other.

Tasks can be carried out individually or in groups (of two or more participants). Prabhu (1987) used no group work at all in his task-based experiment, for reasons that I find rather unconvincing (Prabhu, 1987, pp. 81-83). On the other hand, some teachers use tasks almost exclusively in group work, and that may well be too much of a good thing. In particular, we must take into account that students have different learning styles, and while some may thrive on group work, others may prefer to work on their own. In practice it is likely that a balance of group and individual work, plus a number of tasks in which students choose their way of working, may be the most efficient way of planning task-based courses.[16]

Tasks may be related to one another in different ways. If they are not related at all there are serious problems with the course design. One way of linking tasks is by making the output of one task the input of the next task, creating a chain that culminates in some complex performance or product. Another way is the pre-task/task sequence employed by Prabhu. During the pre-task, operations, language skills, etc., are modelled in a whole-class teacher-learner discussion and demonstration (not unlike the way mathematics problems are dealt with in many classes), and this is then followed by an individual task that parallels the pre-task. Certainly, doing a task more than once is an excellent recommendation. A useful motto is: If it's worth doing, it's worth doing at least twice.

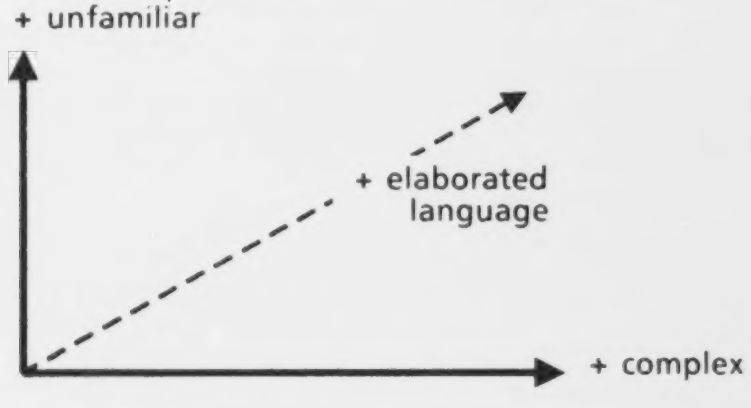
An extension of the pre-task/task design is the scaffolded task-series illustrated above. In fact, Prabhu's pre-task is in a sense a scaffolded task. Further, it is useful to experiment with recording (video or audio) classroom scenes and tasks for subsequent replay, analysis, and discussion. As an example, a practicum class once demonstrated a Japanese lesson that included a cooked Japanese meal, with beautiful table cloths and tableware with Japanese designs. The food was so good, and the class enjoyed the meal so much, that it is doubtful that much attention was paid to the language-learning part of the class. However, we videotaped the session and played it back the next lesson (several interesting scenes repeatedly). All participants felt that the first time, they had enjoyed a great meal but learned little Japanese (only a few isolated words were remembered). But when the tape was replayed, and the scenes were discussed, language learning increased exponentially; I

personally remember several expressions to this day (*arigato, Kitako*!). From the first session, we remembered the food; from the second, the language. I think that this anecdote, uncharacteristic though it may be, holds an important message for classroom learning. It is great to have fun, but L2 learning is efficient only to the extent that cognitive effort is applied to it.

What I have said about tasks applies to content-based language learning equally. Content-based courses can be efficient when the content is challenging and well-taught, and when the students' L2 competence is above a certain threshold. If it isn't, we are merely piling one problem on top of another one, and the students will learn neither content nor language. How we determine this threshold is, I think, a difficult question in need of careful investigation.

When we cannot teach a required chunk of content (e.g., *World History*, or *Introduction to Calculus*) through the medium of L2 because the students' L2 proficiency is simply too low, this does not mean that our L2 lessons should be "contentless," or that the content should be unrelated to what the students are learning in their L1. Far from it. If, however, the linguistic complexity required for a certain L2 task reaches a certain level, it will be necessary to ensure familiarity with the content and the skills needed. Balancing *resources* and *demands* is an important element of task design. Its dynamism is captured in Figure 4, where we can see that the more complex a task is, and the more unfamiliar the students are with the subject matter, the more linguistic elaboration will be required. If we want to push the language forward, it is necessary to start at a level where the students can comfortably handle

FIGURE 4
Familiarity, Complexity, and Linguistic Elaboration



the linguistic requirements for a task. This level may well require a reduction in the complexity of the task itself, or it may require choosing familiar content, or both. Once this level is established, it will be possible to gradually increase the demands. How do we know if the task is too far along on the + elaborated continuum for a particular student or group of students? One frequent indicator is an excessive falling back on the L1 since, if the task is interesting and needs to be done, but the students' L2 resources are unequal to the task, then a logical remedy is to resort to L1. If this resort is denied, a great deal of anxiety or frustration may result.

Monitoring and Action Research

At several places in this paper I have mentioned the importance of monitoring classroom lessons. This monitoring can be done in several ways. First of all, it can be done by self or other. If done by other, this can include students and/or other teachers.

Self-monitoring happens naturally whenever we engage in conscious activity. In fact, being conscious involves monitoring what we are doing. However, it is unclear how usable such monitoring is for assessing the quality of what was done, with an eye to possible improvement for future occasions. While we are teaching, we are very absorbed in what we are doing, and probably have very little mental energy left for careful monitoring. Moreover, our on-the-spot impressions are likely to be quite subjectively colored. Therefore, if we really want useful information about our own actions, we need to take a cool, detached look at them, stepping outside our own shoes for a while, as it were. We can do this by recording our class and listening to/looking at it again later on, perhaps on the car cassette player on the way home, or after the TV news on our VCR. How much such simple re-viewing would influence our subsequent teaching in tangible or intangible ways is very hard to say. It might change our professional life, or it might make no difference at all.

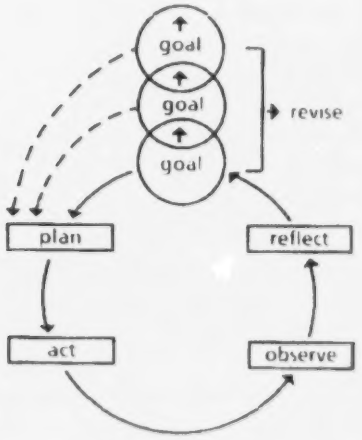
The monitoring can be systematized in several ways. One way that I have used in the past is to record (on audiocassette) every lesson I taught. After listening I would erase (recycle) most of the cassettes, since I didn't feel there was anything worth investigating. However, once in a while there would be a very interesting exchange, conversation, or activity, and I would keep this and transcribe it. The transcription process reveals all sorts of interesting facts and details that were hidden from view before; and it creates insights and raises questions that lead to more focused looking, acting, and listening. This way, a research direction shapes itself.

A particularly fruitful method of monitoring is known as *peer coaching* (Joyce & Showers, 1982). In this format, two teachers arrange

to observe each other's classes, articulating beforehand which particular features they would like to have observed, and conversing about the observations afterwards. This is very useful when we want to try out something new and want to get an impartial (but non-threatening) opinion about how it went. Peer coaching is also used very successfully in teacher education and in supervision and evaluation schemes that employ self- and peer-evaluation rather than evaluation by superiors or administrators.

When such activities of monitoring and discussion (including description and/or analysis, and forward planning) are systematized and carefully documented, we are doing *action research*. Action research can be simply defined as the making of small changes in one's everyday work, and investigating the effect of such changes. It involves a number of cycles of planning, action, observation, and reflection–since every action suggests further (revised) action (Kemmis & McTaggart, 1982; Nixon, 1981). We may feel that we are doing many of these things without wishing to lay claim to the grandiose title of "action research," since this can sound rather official and scientific. It is true that such designations as "teacher as researcher" (Stenhouse, 1975) or "reflective practitioner" (Schon, 1982) refer to ways of working rather than to research activities that bear academic stamps of approval. On the other hand, many teachers who have documented their own working and systematized the steps in the process, have found rich and relevant material for workshops and presentations. The following diagram shows

FIGURE 5
The Action Research Cycle



the way of working of a reflective practitioner, or action research (depending on our goals or ambitions). There are many different visual depictions of action research, all showing the four stages of planning, action, observation, and reflection. However, this particular model emphasizes the goal-driven nature of the work, and the view that our goals become redefined (and refined) as we go along.

Conclusion

In a rather whirlwind tour of the premises, I have tried to show the viability of an L2 classroom learning theory. I have argued that it relies first and foremost on a theory of L2 learning, and this in turn must be based on a general learning theory. I have proposed that awareness, attention, focusing, and vigilance (related terms of increasing focus and energy-intensity) are essential in language learning and that our classroom actions must be designed to encourage and facilitate them.

Authenticity, a notion central to most present-day syllabus design and materials construction, can be shown to be far more complex than the mere provenance of the language, or the superficial similarity of the task to tasks that might happen in non-instructional domains. In particular, *pragmatic*, *pedagogical*, and *existential* authenticity bring new realms of meaning to the word *authentic*.

Interaction, participation, and negotiation are important vehicles for the creation of learning opportunities in the L2 classroom. In order to focus attention efficiently, students must be *intrinsically motivated*. Instrumental and integrative motivation, referring to long-term interests and goals, are beneficial to the extent that they contribute to, or facilitate, intrinsic motivation.

Error correction, an instruction-specific form of repair, is a perennial issue of debate, but I have suggested that it is counter-productive to discuss its pros and cons in isolation from the specific activities in which it occurs. It may be instructive for the reader to study the diagram in Figure 3, and earmark the boxes in which explicit correction of language errors might be appropriate, and other boxes where it would appear to be counter-productive.

The structure of the lesson (or its well-formedness) should be based on a balance between *planning* and *improvisation*, and on an understanding of what is difficult or easy for the students. Grammar exercises and formal explanations are best dealt with outside the lesson (in line with advice from Newmark [1966] and Ellis [1988]), but traditional exercises can be useful in creating points of stability in the lesson, which are necessary for more creative and perhaps less predictable types of work to flourish.

I have discussed several ways in which activities can be *pedagogically scaffolded*, that is, in which systematic guidance

structures are created which are then gradually dismantled (according to the handover principle) as the students indicate they can tackle increasing aspects of roles and communication. Scaffolding is an essential element in a process syllabus, which, in contrast to a macho syllabus, does not pretend to predict in detail what is going to be learned in which way and at what speed. Task-based and content-based teaching and learning can be successful only if they are conducted through a process syllabus, since this will allow for careful monitoring of complexity (as a balance between resources and demands).

Finally, improvements in teaching and learning can occur only if the teaching/learning process is carefully and consistently *monitored*– either by the participants themselves (self-monitoring), or by colleagues (other-monitoring), e.g. by means of *peer coaching*. Every good teacher is a researcher in some sense, or a *reflective practitioner*. But if this everyday searching and researching for questions and answers is systematized and documented, it becomes *action research*.

And so we reach the end of this brief tour. I have tried to be theoretically responsible and practically relevant. In trying to be comprehensive in order to show the feasibility of a *theory of practice*, I may have skated over some surfaces and left some gaps. If, however, the design has remained clear throughout, then we can move forward to create a knowledge base that comes out of the classroom rather than being superimposed on it from the outside.

Acknowledgments

I am grateful for helpful comments from and discussions with a number of people who read earlier drafts of this paper: Kathi Bailey, Kimmarie Cole, Eve Connell, Neil Granoien, Stephanie Stauffer, and Lidia Woytak.

Notes

1. The term L2 learning will subsume, for the sake of convenience, other commonly used terms, such as SLA (second language acquisition), FLL (foreign language learning), SLL (second language learning), and so on. Any differences among these terms will not be presupposed; rather; they will eventually emerge when motivated by the theory.

2. Or "andragogical," in the case of adult learning (see Knowles, 1982). Because of the wider acceptance of the term "pedagogic," I will use it to refer to education in general, regardless of age.

3. Vygotsky defines the zone of proximal development as follows: "It is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers." (1978, p. 86). Putting it another way, "What a child can do with assistance today, she will be able to do by herself tomorrow." (1978, p. 87).

4. I have tried to solve the masculine/feminine pronoun problem by more or less randomly assigning gender to persons, except if I happened to know the gender of a specific person mentioned.

5. Apprehending is akin to cognizing, a term used variously by Kant–from the German *erkennen*–Piaget, Chomsky and others. Apprehension has a narrower meaning in Kant's *Critique of Pure Reason*: There it appears to be a precondition for perception, and is thus closer to my notion of focusing. Arnold's (1960) terms *intuitive appraisal* and *reflective judgment* parallel apprehension and comprehension as used in this paper.

6. I first heard the term uptake in lectures of Dick Allwright, who, if I remember correctly, attributed it to Pit Corder. I am not sure if I do full justice to Pit Corder's original meaning, but in my current scheme of things uptake is close to the traditional meaning of mastery (see also Ellis, 1990, p. 115).

7. A recent book by Andrew Cohen (1990) is a firm step in the right direction. It is very clearly written and practical, being aimed not only at teachers and researchers, but also directly at language learners themselves, although they probably need to be fairly academically oriented and well-educated to profit fully from the book.

8. This figure is actually a primitive or skeletal version of Figure 3, which we will come to later.

9. Kathi Bailey (personal communication) points out that, for this spirit of exploration to exist, the teacher and fellow-students must take care of errors, not to place severe penalties on errors.

10. For an explanation of the transcription conventions used, see van Lier (1988a).

11. I do not make a simple distinction between SL (e.g., ESL in the USA) and FL (e.g., Swahili classes in the USA). Rather, I regard out-of-class exposure as an individual learner variable, where some learners may

use none of abundant exposure that may be available, while others go out of their way to encounter it even where it is scarce. At this level, the SL/FL distinction makes most practical sense, although I acknowledge that in certain settings L2 exposure is harder to obtain (or avoid, as the case may be) than in others. The reader may like to know that S2 is male, S6 female. The extract is taken from van Lier (1988) in slightly edited form.

12. Breen (1985) provides an analysis which is similar in several respects. He discusses authenticity in terms of texts, learners' interpretations, taxes and the social situation of the classroom.

13. The reader may like to know that S2 is male, S6 female. The extract is taken from van Lier (1988) in slightly edited form.

14. Both these orientations are included in the more general category of topic-orientation discussed in van Lier (1988a).

15. The first two of these are called transactional and interactional functions respectively, in Brown & Yule (1983). The textual function is the most difficult one to grasp; it is sometimes called an enabling function; van Lier (1988a, forthcoming).

16. It may also be wise to balance group work with whole-class activities. In particular, the kinds of recommendations made by Egan (1986) for the elementary school, in terms of the centrality of story telling, may well have a great deal of relevance for older students and adults as well.

References

Allen, H.B. & Campbell, R.N. (Eds.) (1972). *Teaching English as a second language: A book of readings*. New York: McGraw-Hill.

Allwright, R. (1984). The importance of interaction in classroom language learning. *Applied Linguistics*, 5, 156-171.

Arnold, M.B. (1960). *Emotion and personality, Vol 1, Psychological aspects*. New York: Columbia University Press.

Beebe, L.M. (1988). Five sociolinguistic approaches to second language acquisition. In L.M. Beebe (Ed.), *Issues in second language acquisition* (pp.43-77). Rowley, MA:

Newbury House.
Bell, A. (1984). Language style as audience design. *Language in Society, 13*, 145-204.
Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge: Cambridge University Press.
Breen, M. (1985). Authenticity in the language classroom. *Applied Linguistics*, 6(1), 60-70.
Brinton, D.M., Snow, M.A., & Wesche, M.B. (1989). *Content-based second language instruction*. Rowley, MA: Newbury House.
Brown, G. & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
Bruner, J. (1983). *Child's talk: Learning to use language*. New York: Norton.
Candlin, C.N. & Murphy, D.F. (Eds.) (1987). *Language learning tasks*. Lancaster Practical Papers in English Language Education, Vol.7. Englewood Cliffs, NJ: Prentice-Hall.
Cathcart, R.L. (1989). Authentic discourse and the survival English curriculum. *TESOL Quarterly, 23* (1), 105-126.
Cazden, C.B. (1983). Peekaboo as an instructional model: Discourse development at school and at home. In B. Bain (Ed.) *The sociogenesis of language and human conduct: A multi-disciplinary book of readings* (pp. 33-58). New York: Plenum Press.
Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
Clarke, M.A. (1983). The scope of approach, the importance of method, and the nature of techniques. In J.E. Alatis, H.H. Stern, & P. Strevens (Eds.), Georgetown University Roundtable on Language and Linguistics (pp.106-115). Washington, DC: Georgetown University Press.
Cohen, A.D. (1990). *Language learning: Insights for learners, teachers, and researchers*. Rowley, MA: Newbury House.
Crookes, G. & Schmidt, R. (1989). Motivation: Reopening the research agenda. *University of Hawai'i Working Papers in ESL*, 8 (1), 217-256.
Danesi, M. (1988). Neurological bimodality and theories of language teaching. *Studies in Second Language Acquisition*, 10, 13-31.
Deci, E.L. (1975). *Intrinsic motivation*. New York: Plenum.
Deci, E.L. & Ryan, R.M. (1985). *Intrinsic motivation and*

self-determination in human behavior. New York: Plenum.

Dewey, J. (1904). The relation of theory to practice in education. *Third Yearbook*, Part I (pp. 9-30). Bloomington, IL: National Society for the Study of Education.

Egan, K. (1986). *Teaching as story telling*. Chicago: The University of Chicago Press.

Ellis, R. (1984). *Classroom second language development*. Oxford: Pergamon Press.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1988). The role of practice in classroom language learning. In G. Kasper (Ed.), *Revue de l'AILA/AILA Review 5*, 2039. Gent: Association Internationale de Linguistique Appliquee.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Basil Blackwell.

Fiske, D.W. & Shweder, R.A. (1986). *Metatheory in social science: Pluralisms and subjectivities*. Chicago: The University of Chicago Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, R.C. (1985). *Social psychology and second language learning*. London: Edward Arnold.

Gardner, R.C. 1991). Student variables in language learning. *Applied Language Learning, 2(1), 1-28*.

Habermas, J. (1987). *The theory of communicative action. Vol. 2, Lifeworld and system: A critique of functionalist reason*. Boston: Beacon Press.

Hawkins, E. (1987). *Language Awareness*. Cambridge: Cambridge University Press.

Joyce, B.R. & Showers, B. (1982). The coaching of teaching. *Educational Leadership*, 40, 4-10.

Kemmis, S. & McTaggart, A. (1982). The action research planner. Victoria, Australia: Deakin University Press.

Knowles, M. (1984). *The adult learner: A neglected species*. Houston: Gulf Publications

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Krashen, S. (1985). *The input hypothesis*. New York: Longman.

Krashen, S. (1990). Plenary address. *Reading, writing,*

form and content. Georgetown University Round Table on *Languages and Linguistics*.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77-99). San Diego: College-Hill Press.

Mackey, W.F. (1965). *Language teaching analysis*. London: Longman.

Magee, B. (1973). *Popper*. London: Fontana.

Mohan, B. (1986). *Language and content*. Reading, MA: Addison-Wesley.

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Newmark, L. (1966). How not to interfere with language learning. IJAL, *32*, 77-83. Reprinted in H.B. Allen & R.N. Campbell (Eds.) (1972). *Teaching English as a second language: A book of readings*. (2nd ed.) (pp.37-42). New York: McGraw-Hill.

Newmeyer, F.J. & Weinberger, S.H. (1988). The ontogenesis of the field of second language learning research. In S. Flynn & W. O'Neill (Eds.), *Linguistic theory in second language acquisition* (pp.34-45). Dordrecht: Kluwer Academic Publishers.

Nixon, J. (Ed.) (1981). *A teacher's guide to action research*. London: Grant McIntyre.

O'Malley, J. M. & Chamot, A.U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oskarsson, M. (1973). Assessing the relative effectiveness of two methods of teaching to adults. *IRAL, 11* (3), 251-262.

Peirce, C.S. (1955). Abduction and induction. In J. Buchler (Ed.), *Philosophical writings of Peirce* (pp.150-156). New York: Dover Publications.

Peters, A. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.

Piaget, J. (1975). *The grasp of consciousness: Action and concept in the young child*. Cambridge, MA: Harvard University Press.

Prabhu, N.S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Prabhu, N.S. (1990). There is no best method - why?

TESOL Quarterly, 24 (2), 161-176.
Price, H.H. (1969). *Thinking and experience*. London: Hutchinson University Library.
Rudduck, J. & Hopkins, D. (Eds.) (1985). *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse*. London: Heinemann.
Schegloff, E.A., Jefferson, G. & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*, 361-382.
Schmidt, R.W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.
Schmidt, R.W. & Frota, S.N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R.R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp .237-326). Rowley, MA: Newbury House..
Schon, D. (1977). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
Skehan, P. (1989). *Individual differences in second-language learning. London*: Edward Arnold.
Spinoza, B. (1910). De intellectus emendatione. London: J.M. Dent & Sons.
Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann.
Sternberg, R.J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Viking.
Stubbs, M. (1983). *Discourse analysis*. Oxford: Basil Blackwell.
Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal/Revue TESL du Canada, 6*(1), 68-83.
van Lier, L. (1988a). *The classroom and the language learner: Ethnography and second-language classroom research*. London: Longman.
van Lier, L. (1988b). What's wrong with classroom talk. *Prospect*.
van Lier, L. (forthcoming). *Language awareness: Putting language back into language teaching*.
Vygotsky, L.S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
Widdowson, H.G. (1979). *Teaching language as communication*. Oxford: Oxford University Press.
Wittgenstein, L. (1958). *The blue and brown books*. Oxford: Basil Blackwell.

Author

LEO VAN LIER, Associate Professor of Educational Linguistics,
Monterey Institute of International Studies,
425 Van Buren Street, Monterey, CA 93940.
Specializations: classroom language learning, discourse analysis, language in education.

Leading the field in sociology and the related social sciences:

sociological abstracts (sa)

and

Social Planning / Policy & Development Abstracts (SOPODA)

Our subject specialists track the broad spectrum of theoretical and applied sociology from the more than 1,800 discipline-specific and related journals published in North America, Europe, Asia, Africa, Australia, and South America.

sa and **SOPODA** each offer you in-depth abstracts and precise indexing of timely journal articles and books, enhanced dissertation listings, and a bibliography of book reviews from the journals screened.

sa and **SOPODA** are available together on the **sociofile** CD-ROM and are hosted online by BRS, DATA-STAR, DIALOG, and DIMDI. Hardcopy subscriptions can be ordered from the address below.

The **sa** and **SOPODA** information products are supported by:

- Database-specific user manuals
- The latest journal coverage list
- The **sociofile** Quick Reference Guide and User's Handbook
- The **Thesaurus of Sociological Indexing Terms**
- Your Guide to Searching **sa** using a Personal Computer
- A professional workshop program

The **sa** family of databases — your fast track to the information you need, in the format you want.

sociological abstracts, inc.

p.o. box 22206 • san diego, ca 92192-0206
phone (619) 695-8803 / FAX (619) 695-0416 / Help Desk (800) 752-3945

Applied Language Learning
1991, Vol. 2, No. 1, pp.71-113

Measurement and Research Implications of Spolsky's Conditions for Second Language Learning

John L. D. Clark
Defense Language Institute

Francis E. O'Mara
PRC, Inc.

Bernard Spolsky's recently-published *Conditions for Second Language Learning* (Spolsky, 1989) is a substantive contribution to the second/foreign language learning field. In keeping with and expanding upon several of his earlier publications on language learning and language learning measurement isues, Spolsky takes, in the referenced volume, an extremely holistic perspective on language learning, viewing it as a process in which a large number of social, cultural, and other influences regularly and inevitably operate, together with those usually discussed in the language teaching literature, to influence and ultimately determine both the nature and extent of the learning outcome. The primary purpose of the Spolsky volume is to define and elaborate on 74 such factors or "conditions," and to describe the manner in and degree to which these conditions may be considered as contributing, both individually and in combination, to the language learning process.

Among the 74 conditions for language learning which he posits, Spolsky identifies 31 as "necessary" conditions, that is, conditions "without which [language] learning is impossible" (p. 14). The remaining conditions are viewed as being "typically" involved in the production of a given linguistic outcome, although not invariably required in the attainment of that outcome. A third descriptor, "graded," is applied by Spolsky to many of his conditions (both necessary and typical) to indicate that, for these conditions, "there is a relation between the amount or extent to which a condition is met and the nature of the outcome" (p. 14).

To facilitate later discussion of the measurement considerations at issue, it is useful to categorize the learning conditions identified by Spolsky under four major headings (present writers' classifications). First, variables associated with the particular *language* being learned, especially as these interact with the linguistic features of the learner's mother tongue, are considered by Spolsky to exert a major influence on the learning process. Spolsky describes this overall effect as follows:

> Condition 34. Language Distance condition (necessary, graded): The closer two languages are to each other genetically and typologically, the quicker a speaker of one will learn the other. [117][1]

At a more detailed level, other language-related characteristics are seen to come into play in determining which aspects of the language will be learned first and/or more thoroughly, for example:

> Condition 37. Markedness Differential condition (necessary, graded): Marked features [which are not also marked in native language] are more difficult to learn than unmarked. [122]

Second, in addition to the inherent formal characteristics of the target language itself, the degree of cultural, political, religious, or other types of status accorded that language within the *social context* in which the language learning take place constitute, in Spolsky's model, an important set of conditioning variables. Social context variables include, for example:

> Condition 42. Number of Speakers condition (typical, graded): The number of people who speak a language as a first or second language influences the desire of others to learn it. [133]
>
> Condition 45. Official Use condition (typical, graded): Prefer to teach or learn a language which is officially used or recognized. [136]
>
> Condition 47. Great Tradition condition (typical, graded): Prefer to teach or learn a language which has a desirable Great Tradition (including a religion) associated with it. [136]

[1] Throughout this paper, condition definitions or other substantial quotations from the Spolsky text are shown in indented blocks, with page references in square brackets.

A third and extremely powerful set of conditions involves the variety of skills, capacities, relevant knowledge, attitudes toward and motivation for language learning, and other personal characteristics which the *language learner* brings to the learning setting. Within this wide range of learner-specific variables, a subset is concerned with individual physical characteristics over which neither the learner nor the instructional process may be expected to have any appreciable control, including:

> Condition 22. Physiological Normality condition (necessary): Any physiological or biological limitations that block the learning of a first language will similarly block the learning of a second language. [89]

> Condition 28. Sound Discrimination condition (necessary, graded): The better a learner can discriminate between the sounds of a language and recognize the constituent parts, the more successful his or her learning of speaking and understanding a second language will be. [106]

Another subset of individual learner variables is associated with generalized, longer term cognitive capacities developed through formal schooling or other relevant life experiences. Although not so unalterably "fixed" within a given individual as the preceding category of physiological variables, these general cognitive variables are not usually amenable to meaningful intervention or change within the context of a particular language learning situation per se. Examples of general cognitive variables include:

> Condition 24. Abstract Skills condition (typical, graded): Formal classroom learning of a second language is favored by the development of skills of abstraction and analysis. [97]

> Condition 30. Grammatical Sensitivity condition (necessary, graded): Beyond the necessary minimum ability to 'derive a grammar' implicitly, the better a learner's ability to recognize constituents and develop or understand generalizations about recombination and meaning (whether from explicit or implicit generalizations, in whatever forms), the faster he or she will develop control of the grammatical (and pragmatic) structure of a second language. [106]

Yet another subset of learner variables includes those which can reasonably be addressed within the context of the language program itself, through suitable analysis and adaptation in the manner in which the learning material is presented to the learner, based on the individual's particular learning style and modality preferences; as well as by dealing effectively with affective aspects of the learning process. Examples include:

> Condition 33. Second Language Learning Anxiety condition (typical, graded): Some learners, typically those with low initial proficiency, low motivation, and high general anxiety, develop levels of anxiety in learning and using a second language that interfere with the learning. [115]

> Condition 31. Learning Style Preference condition (typical, graded): Learners vary (both individually and according to such characteristics as age, level, and cultural origin) in their preference for learning style (visual, auditory, kinaesthetic, and tactile) and mode (group or individual); as a result, learning is best when the learning opportunity matches the learner's preference. [110]

Fourth and finally among the major learning conditions categories are the diverse *opportunities for learning* the target language that are present and operative in the learning environment, either in formal, structured settings typical of academic language instruction and/or in informal, naturalistic settings characteristic of untutored or "street" learning. Examples of opportunities-for-learning conditions that apply to both formal and informal language learning contexts include:

> Condition 51. Exposure condition (necessary, graded): The more time spent learning any aspect of a second language, the more will be learned. [148]

> Condition 57. Opportunity for Analysis condition (necessary, graded): Learning a language involves an opportunity to analyze it, consciously or unconsciously, into its constituent parts. [167]

> Condition 60. Opportunity for Matching condition (necessary, graded): Learning a language involves an opportunity for the learner to match his or her own

knowledge with that of native speakers or other targets. [169]

Other opportunities-for-learning conditions may be considered characteristic of only formal learning situations, for example:

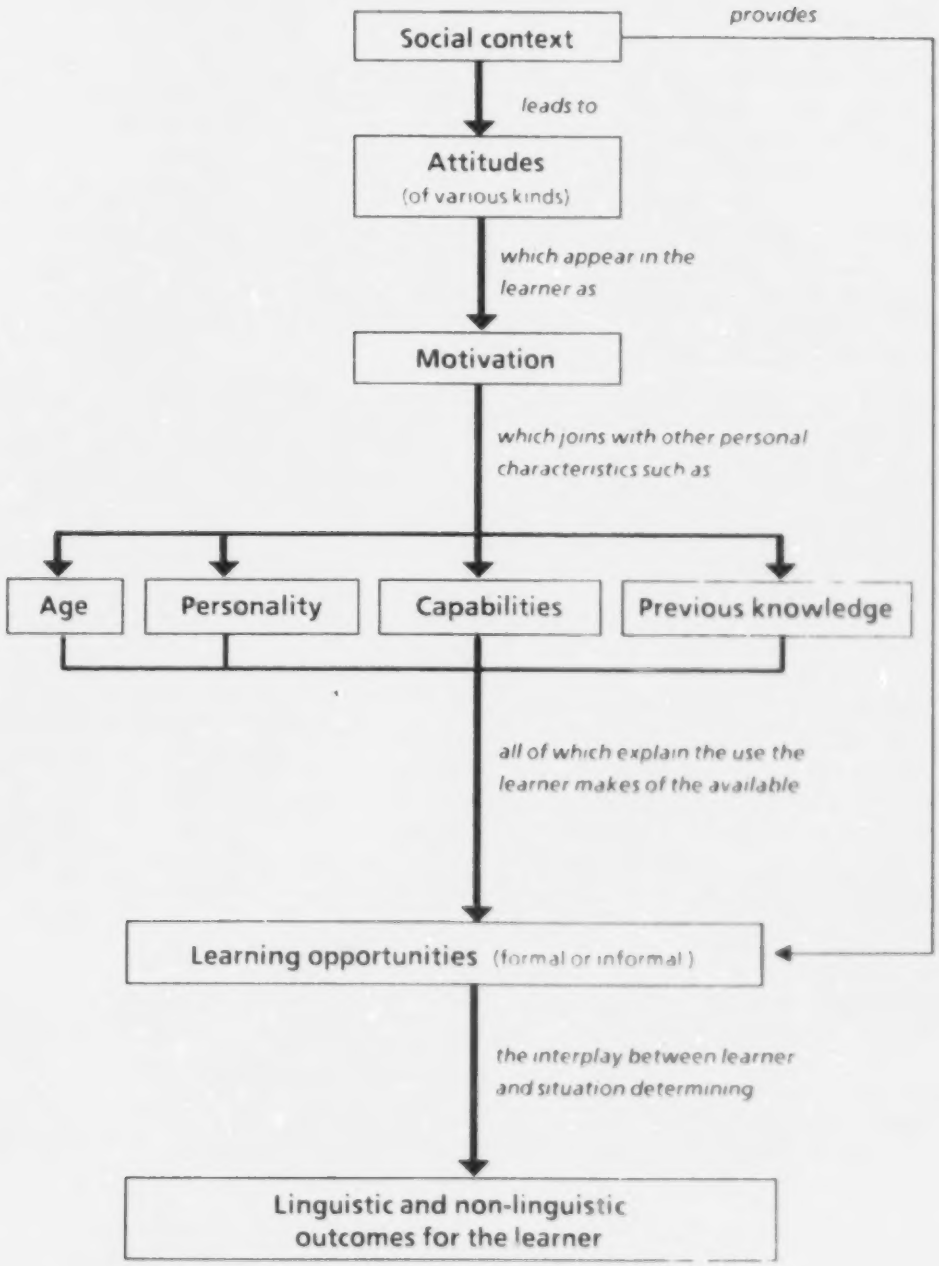
> Condition 68. Classroom condition (typical of formal learning, graded): The learning takes place in a closed physical space, a single classroom. [172]
>
> Condition 70. Simplified Language condition (typical of formal learning, graded): The language is simplified and controlled. [172]
>
> Condition 66. Teacher Model condition (typical of formal learning, graded): Only one speaker (the teacher) is fluent; the majority in the environment (classroom) are not. [172]

Learning conditions associated with informal learning situations are, by contrast, typically quite different from those encountered in the structured classroom setting. Examples of informal learning conditions include:

> Condition 67. Open Area condition (typical of natural learning, graded): The learning takes place in the open or in unconstrained areas. [172]
>
> Condition 69. Uncontrolled Language condition (typical of natural learning, graded): The language is normal and uncontrolled. [172]
>
> Condition 65. Fluent Speakers condition (typical of natural learning, graded): Many speakers in the environment are fluent and native. [172]

The posited functional interrelationships of the various types of learning conditions described by Spolsky are represented schematically by him in the following figure, which he titles and narratively describes as "a model of second language learning" (p.28).

FIGURE 1
A Model of Second Language Learning



The power and potential utility of the Spolsky model as a framework for characterizing and analyzing any of a wide variety of language learning situations lies in its generality and inclusiveness; even the most assiduous reader of the Spolsky volume would be hard pressed to identify, for any real-life language learning context, either formal or informal, any componential variable that is not explicitly or implicitly included within one or another of the 74 learning conditions which he defines.

Spolsky is equally synoptic in his treatment of the possible *linguistic outcomes* at issue across the wide variety of potential language learning undertakings in both formal and informal contexts. In this regard, he refers to three types of assertions that can be made with respect to an individual's competencies or abilities in a second language, differentially characterizing these assertions as the *structural* claim, the *functional* claim, and the *general proficiency* claim. Each of these claims, according to Spolsky, "derives from a different theoretical concept of the nature of knowledge of a language, and each has different empirical consequences." (pp. 59-60) The *structural* claim is associated with the view that "knowledge" or "ability" in the language is reflected in the extent of the learner's knowledge of the grammatical structures and lexicon of the target language. Language tests developed in accordance with this perspective tend to emphasize discrete-item formats addressing the learner's ability to "handle" each of a variety of structural elements (e.g., the past tense forms of certain verbs) or types/categories of lexicon (e.g., parts of the body, public transportation terms). The *functional* claim coincides with the general approach to language learning that emphasizes the learner's ability to carry out specified communicative tasks using the language. Tests developed on the basis of this model focus on assessing the learner's ability to satisfy real-life communicative needs. Munby's (1978) detailed analyses of the linguistic tasks performed by particular individuals in highly specified language-use settings (such as those of a waiter in a particular type of restaurant) are exemplary of the functional approach to teaching and testing issues. The *general proficiency* claim is based on the assumption that, although it may as a practical matter be impossible to specify the entire corpus of structures, lexicon, and functional capabilities required for full (native-like) performance in the language, it is nonetheless possible to speak meaningfully of lesser or greater degrees of general "proficiency" on the part of a given language learner, as evidenced by lesser or greater control of structure, lexicon, and functions as measured by tests targeting these elements, or by a "special kind of functional test that taps the generalized ability more directly." (p. 72) The Interagency Language Roundtable scale and associated interview procedure, as well as the derivative ACTFL guidelines and testing approach, are characterized by Spolsky as falling within the "general proficiency"

category, although with his important caveat that these scales are for the most part based on a somewhat restricted corpus of functional language needs: those at issue in polite social interaction, tourism or temporary in-country residence, and job performance in various representational contexts abroad.

The complex issue of developing testing formats and test content capable of validly assessing those learner attainments produced through the operation and interaction of language, social context, language learner, and opportunities for learning variables within a given teaching/learning setting is summarized by Spolsky in a multifaceted Linguistic Outcome condition, which he describes as follows:

> Condition 20. Linguistic Outcome condition (typical, graded): Prefer to say that someone knows a second language if one or more criteria (to be specified) are met. The criteria are specifiable:
>
> (a) as underlying knowledge or skills
> (b) analyzed or unanalyzed
> (c) implicit or explicit
> (d) of individual structural items (sounds, lexical items, grammatical structures)
> (e) which integrate into larger units
> (f) such as functional skills
> (g) for specified purposes
> (h) or as overall proficiency
> (i) productive or receptive
> (j) with a specified degree of accuracy
> (k) with a specified degree of fluency
> (l) and with a specified approximation to native speaker usage
> (m) of one or more specified varieties of language [80]

The complexity of the learning outcome measurement picture drawn by Spolsky is by no means surprising in light of the highly open architecture of the other components of his learning model. Indeed, this degree of flexibility in specifying the nature of the testing process and the content of the associated test instruments is absolutely necessary in order to provide sufficient latitude for accurate and meaningful evaluation of the large variety of alternative learning outcomes that can be identified and embraced within the Spolsky model.

Implications of the Spolsky Model for Research on Language Learning / Teaching

Although not explicitly characterized as such by Spolsky, the great majority of his learning conditions can readily and very productively be categorized as falling into one of three operational categories of potential research variables: *input*, *process*, and *output* (Clark, 1988a). Input variables are those associated with the "raw material" on which the experimental process or treatment is performed, and comprising, in the language-learning context, the student or other learner and the multifaceted physiological, experiential, and cognitive characteristics which he or she brings to the learning task. Process variables are those involved in the instructional/learning activity itself, including not only formally-arranged learning contingencies (e.g., specific activities within a planned academic curriculum) but also any and all adventitious, nonstructured learning opportunities available in the day-to-day living situation and societal context in which the learner is immersed. Output variables represent the "products" of the learning activity, as defined and quantified by the particular assessment instruments or procedures administered on completion of the identified learning activity.

The most significant contribution of the Spolsky model to the cause of valid and meaningful research on language learning/teaching is to greatly broaden the field's perspective as to the number and variety of constituent elements that may be considered to be in operation, and in complex interaction, in any language learning--or associated research--setting. Directly related to this major conceptual contribution is the significant operational challenge of developing, and broadly adopting within the profession, measurement instruments, research approaches, and analysis procedures that will make it possible to take into formal account as many as possible of the input, process, and output variables at issue in any given language study. In this regard, at least five large-scale, coordinated efforts on the part of the language learning research community would appear to be required, specifically, to:

(1) identify or develop valid, reliable, and practical tests or other instruments to operationally define and quantify the intended lingusitic or other "output" results of language learning activities, so as to be able to standardize learning outcome measures as much as possible across differing learning studies and research contexts;

(2) identify or develop standardized, broadly accepted and endorsed measures of (or, at the least, highly objective descriptors for) the major variables at issue within the "input" variable category, including the *language*, *social context*, and *language learner* condition sets elaborated by Spolsky;

(3) identify or develop adequate measures or descriptors for at least the most salient "process" variables at issue within Spolsky's *opportunities for learning* conditions;

(4) identify or develop appropriate statistical and other analytical procedures to permit informed interpretation of the observed relationships among the obtained output data and the broad constellation of input variables and process variables at operation in any given language learning setting;

(5) develop and implement organizational mechanisms within the second language teaching and research community to encourage and facilitate increased standardization and interoperability of research studies in terms of: (a) the tests or other instruments used for output performance measurement, as well as those used for input- and process-variable measurement (to the extent that quantifiable data can be gathered in these latter two areas); (b) the protocols for or verbal descriptors of input and process variables not reasonably amenable to instrument-based quantification; and (c) the inclusion, in the final research report, of all necessary information to properly situate the study within the larger context of an overall research framework such as that provided by the Spolsky model.

Although it is well beyond the scope of this article to attempt to describe each of the types of instrumentation needed to quantify or otherwise objectively characterize all of the input, process, and output variables at issue in the Spolsky model, several examples of possible instrumentation/documentation procedures within each of these three categories are provided, with a view to typifying the kinds of tests or other data gathering procedures that might be identified and codified in subsequent, more intensive efforts to fully operationalize the Spolsky model. Following this, the development of statistical analysis procedures capable of simultaneously accounting for the input, process, and output variables at issue in a given language learning context is discussed, and suggestions made for more thoroughly standardizing language learning research and reporting procedures so as to take maximum advantage of communally available or to-become-available instruments, descriptive protocols, and analytic techniques.

Output Variables

As discussed at greater length in a recent separate article (Clark, 1988a), the proper design and development of any language learning/teaching activity, of whatever type or scope, depends crucially on the clear and detailed specification of the intended linguistic outcome of this activity. By the same token, any valid attempt to determine the yield or efficacy of a given learning/teaching undertaking is crucially dependent on having available a clear and detailed description of the intended outcome goals, together with measurement instruments or other structured assessment procedures that embody and operationally define these goals. Fortunately, at least for the language testing community, among the three components of the input-process-output model, output variables are probably the most amenable to objective description and quantification, notwithstanding the existence of substantial and continuing measurement challenges in this area as well.

With respect to "output" measurement, it would be both inappropriate and counterproductive to attempt to identify or advocate a single, one-measurement-type-fits-all approach to assessing language teaching/learning outcomes across the very large number of potential outcomes subsumed by the Spolsky model. Spolsky's own categorization of possible output measures into tests addressing, respectively, *structural*, *functional*, and *general proficiency* aspects of performance is of some usefulness as a point of departure, and coincides reasonably well with the senior author's recently proposed classification of possible testing purposes into two major categories consisting of (a) the *provision of instructional feedback* intended to inform and facilitate the language learning process, and (b) the *assessment of competence* to carry out various real-life activities using the language (Clark, 1989). Structurally-oriented testing, as characterized by Spolsky, would appear to fall rather clearly within the instructional feedback category, while both functional and general proficiency testing may be seen as capable of serving either for instructional feedback (albeit of a quite summative nature) or for competence assessment in non-instructionally oriented measurement contexts (e.g., in connection with employment applications, verification of "fitness" to undertake certain language-related assignments, and so forth).

With regard to the design and use of both instructional feedback (generally, "structural") tests and competence (generally, "functional" or "proficiency") tests, a major consideration is the essentially inverse relationship between the *diagnostic precision* of the test and its *linguistic authenticity*. For example, a test aimed at determining the examinee's ability to aurally distinguish present-, past-, and future-

tense structures within a series of spoken sentences, while highly diagnostic with respect to the particular structural elements at issue, would typically result in "made-up" texts not ordinarily encountered in authentic language-use contexts. On the other hand, a highly authentic listening comprehension test comprised of texts drawn directly and without adaptation from genuine real-life sources (e.g., radio or television broadcasts) would almost inevitably include a considerable amount of linguistic material extraneous to, and potentially confounded with, assessment of the examinee's "tense-discrimination" ability per se.

For these and other reasons as discussed more fully in a separate paper (Clark, 1989), it may be suggested that in order to best serve the differing measurement goals at issue in diagnostically-oriented "structural" assessment on the one hand and "function-" or "proficiency-" oriented testing on the other, the most appropriate and productive development strategy will probably be to design and elaborate each type of test separately and in accordance with its own intrinsic measurement rationale, rather than to attempt to produce a hybrid instrument that would in all likelihood not serve either informational function as well as tests devoted to a single measurement purpose.

With regard to "structural" testing, it is unfortunately not possible to be highly sanguine as to the prospects of making standardized, communally usable instruments routinely available for use in this area. Although some attempts have been made to produce diagnostically-oriented, discrete-item tests that would be uniformly applicable across a variety of learning settings (see, for example, the tests developed for use in the University of Pennsylvania/Center for Applied Linguistics' recent Language Attrition Study project [Lambert & Stansfield, 1988]), the problem of identifying a relatively small and universally "known" test lexicon that could be used in assessing examinee control of other-than-lexical (e.g., structural, phonological) elements without introducing lexical familiarity itself as a confounding variable has proved highly resistant to solution. A second factor militating against the use of discrete-element tests as a common measure across diverse learning settings is the test scoring and reporting burden associated with this type of test, which, in effect, provides a series of dichotomous (i.e., "knows/controls"; "does not know/control") values for each of the numerous individual elements tested, as opposed to a single "global" performance score.

With respect to "functional" or "general proficiency" testing, the profession is fortunate in having available to it, or increasingly under development, a number of instruments which can be used with some degree of justification and effectiveness as "common metrics" in assessing the output of a wide variety of language learning programs and activities. The Educational Testing Service has developed

proficiency-oriented listening comprehension and reading tests in Russian, and the Center for Applied Linguistics has published a listening/reading battery in Chinese, with similar tests in preparation in several additional languages. With regard to speaking proficiency, the American Council on the Teaching of Foreign Languages' training and certification program for oral proficiency testers offers a reasonable means of objectifying and standardizing assessment activities in this skill modality. In addition, the so-called "semi-direct" speaking tests (Clark, 1979), which deliver the test stimuli via booklet and tape recording, with the examinee's responses recorded for later evaluation, provide a reasonable and highly standardized alternative to face-to-face testing. Tests of this type are currently available for Chinese and Portuguese through the Center for Applied Linguistics, with additional tests under way in other languages. Although somewhat less face and procedurally valid than the direct interview, the semi-direct tests have been found to correlate highly with contemporaneously administered face-to-face interviews (Cartier, 1980; Clark, 1988b; Stansfield et al., 1990), and may be considered a reasonable alternative approach to measuring speaking ability when direct interviewing is not economically or administratively feasible.

The availability to, and for the most part positive reception by, the language teaching community of these and other tests of general language ability should not be allowed to obscure or minimize the lively debate among language testers themselves over the possibility of devising a single measure of "global" language performance that is not itself subject to criticism as being conceptually or operationally dependent on a particular type of instruction and/or particular sequence of language learning activities. (See, in this regard, Bachman and Savignon [1986], Clark and Clifford [1987], Clark and Lett [1988], Lado [1978], Lantolf and Frawley [1985], and Lowe [1986]). The point to be urged here is that notwithstanding the useful and indeed essential debate surrounding proficiency testing issues in general, and pending further conceptual and operational refinements in the measurement techniques employed, the language profession would be very well served by agreeing to broadly adopt and use in its research efforts currently available general proficiency tests, so as to be able to uniformly define and analytically control for at least the output component of the input-process-output paradigm in conducting and reporting research studies across a variety of languages and learning situations. In the absence of adequate standardization of the output measures employed, the likelihood of obtaining valid and broadly extrapolatable empirical information on the contribution of input and/or process variables to second language learning is considerably reduced.

Input Variables

Once the intended language-performance output of a given teaching/learning undertaking has been determined and suitable measurement instruments identified and developed, empirical attention can appropriately be focused on the detailed description and, to the extent possible, quantification of the large variety of output-affecting conditions existing within the individual language learners themselves, who constitute the "input" upon which the teaching/learning process is intended to operate. Within the last several years, considerable progress has been made in operationally defining, as well as developing assessment instruments for, several of the presumed most critical learner variables, and the point is rapidly approaching at which second language researchers should have available many of the tools needed to quantify and make solid interpretative sense of the "input" component of the three-element paradigm.

Within the input variable domain, it is possible to identify several major subsets of variables as briefly described below, together with an indication of the types of instruments or procedures considered appropriate to quantifying these variables for research and analysis purposes.

Physiological variables. (E.g., Spolsky's Condition 22: Physiological Normality.) Language learning studies reported in the literature rarely if ever test explicitly for possible physiological abnormalities on the part of the learners. Although it would not be cost-effective or particularly useful to conduct and report detailed physiological "work-ups" on each of the participants in, for example, a large-scale, school-based learning study, it would appear reasonable and administratively practicable, even in such contexts, to verify through existing records that none of the subjects have clinically significant auditory, speech, or visual acuity deficits or other condition reasonably anticipated to affect the language learning process (e.g., dyslexia). Those studies addressing a particular language skill modality or modality sub-aspect in which physiological factors would be expected to play an especially salient role (e.g., accuracy of tone discrimination as related to listening comprehension in Chinese) would be expected to measure and report in greater detail each subject's baseline performance on relevant standardized physiological tests--in this example, overall hearing acuity, pitch discrimination accuracy, and so forth. Age and gender of the subjects would also of course be important "input" variables, particularly for studies addressing aspects of language learning that are hypothesized or have been found to be influenced by

these variables (e.g., age as a conditioning factor in the acquisition of native-like pronunciation).

General cognitive variables. (E.g., Condition 27: Intelligence; Condition 29: Memory.) Given the extensive research literature documenting strong and consistent relationships between general intelligence and language learning effectiveness, it is regrettable that standardized measures of these variables are not administered and scores reported as a matter of course in both small-scale and larger-scope learning studies. Particularly in school-based studies, where student academic files are generally accessible for legitimate research purposes, it should be readily possible to routinely gather standardized intelligence test scores as part of the routine input-data variable set.

Language learning aptitude. (Condition 50: Aptitude) This category includes essentially invariant mental characteristics of the learner (other than "general cognitive" traits) that have been found to correlate positively with the speed or thoroughness of second language learning in given learning contexts. Although an absolutely clear and unambiguous distinction between general cognitive ability and language-learning aptitude has yet to be drawn at both the conceptual and operational (test instrumentation) levels, it is nonetheless heuristically useful to posit the existence of the latter category of learner-specific traits and attempt to measure their relationship to language learning outcomes. Pending the development of other prospective measurement devices or procedures in this area, the *Modern Language Aptitude Test* (Carroll & Sapon, 1959) and the *Pimsleur Language Aptitude Battery* (Pimsleur, 1966) are both readily available and merit strong consideration as routinely administered "input" measures in language learning studies generally.

Psychological variables. Included in this category are learner-specific characteristics discussed in the literature under either the rubric of "attitude (toward)" or "motivation (for)" language learning. Psychological variables may be distinguished from physiological, general cognitive, and language learning aptitude factors in that psychological variables are appreciably more amenable to change as a function of the learner's specific experiences within the general cultural or social environment and/or as an intentional or unplanned consequence of the language learning process itself. A variety of potentially useful attitudinal and motivational measures have been developed over the past twenty years (see Gardner and Lambert [1972], as well as the Gardner article in the present volume), although they are not as readily available as the commercially distributed general intelligence and language aptitude tests. A further consideration that complicates the effective, highly standardized measurement of attitudinal/motivational variables is that it is usually necessary, for measurement validity purposes, to tailor the content and specific

wording of both types of instruments to the age and experiential background of the respondents, as well as to the particular learning or teaching context at issue in a given study.

In addition to attitudinal and motivational factors, a growing body of research suggests that individual personality characteristics may play a substantial role in language learning outcomes. For example, Madeline Ehrman's recent use of the *Myers-Briggs Type Indicator* (MBTI) in learning studies conducted at the Foreign Service Institute (Ehrman, 1990) points to the utility of the MBTI as a classificatory device for both the student (on the input-variable side of the learning model) and the instructor (as one component of the process variable data set). One implication of Ehrman's work which merits further carefully controlled experimentation is that both instructors and students may need to be appropriately "matched" to one another with respect to MBTI-measured personality characteristics in order to optimize the effectiveness of their interaction in the teaching-learning process.

Prior knowledge of target language or other second language. Although not explicitly included in his overall list of conditions, the nature and extent of the learner's prior exposure to the target language (and/or other non-native languages) is mentioned by Spolsky as an important "input" variable in language learning analyses (p. 27). From a measurement perspective, some appropriate and reliable assessment of the learner's initial state of performance in the target language (i.e., just prior to initiation of the learning activities under examination) would appear to be required. Depending on the scope and information-gathering goals of the study, this assessment could be at the level of measured general proficiency in the language or could focus more narrowly on the command or lack of command of specific target language structures or other discrete linguistic elements. Direct measurement of learner performance just prior to beginning the learning program would of course be the psychometrically preferred approach. However, for situations in which this is not administratively or otherwise feasible, learner self-reports, if properly structured and elicited, may be a reasonable alternative approach. Recent studies comparing self-assessment procedures to direct competence measures have shown fairly high correlations in several instances, particularly when the learner is asked to estimate the degree of success with which he or she is able to accomplish each of a series of specified functional tasks (as with the so-called "can-do" scale developed by Clark [1981] and related measures employed by Gardner et al. [1987] and Hilton et al. [1985]), as opposed to making more global judgments about his or her overall "proficiency" in the language.

The possibility of direct empirical measurement of the learner's competence in one or more previously-studied or informally learned languages other than the target language of interest must be considered,

as a practical matter, low or nonexistent in most real-life research settings. Questionnaire-based data gathering should, however, be quite feasible in most instances, and would be particularly called for in situations in which an appreciable proportion of the learners would be expected to have had instruction in or contact with other-than-target languages in the course of their formal schooling or other life experience.

Learner style and learning preference variables. Considerable progress has been made within the past few years in operationally defining and quantifying differences in the ways in which individual learners characteristically go about the business of learning a language in both formal and informal contexts. The most ambitious instrumentation efforts to date have been those of Rebecca Oxford and her associates, culminating in the publication of the *Strategy Inventory for Language Learning* (SILL), a 121-item multiple-choice questionnaire based on a detailed taxonomy of possible student-initiated learning strategies for learning a second or foreign language, ranging from memorizing word lists to seeking out native speakers for conversational practice (see Oxford and Nyikos, 1989; Oxford, 1990). Factor analyses of SILL data obtained on a large sample of intensive course students at the Defense Language Institute produced five main factors, characterized by Oxford as formal rule-related practice; functional practice; use of resourceful, independent strategies; general study strategies; and conversational input elicitation. To the extent that preferred approaches to the language learning task are found to vary across individuals, it would be important to be able to characterize these differences in objective terms through use of the SILL or other standardized instruments for routine input-variable data gathering in language learning studies.

Process Variables

Probably the most challenging component of the input-process-output model with respect to objective quantification and measurement is the language teaching/learning process itself. A major consideration, and corresponding source of difficulty in this regard, is the extremely broad range of applicability of the term, "process." Although not explicitly described by Spolsky as such, the total conceptual scope of the numerous process variables cited in the *Conditions* volume would appear to include any and all external situations, stimuli, or other influences encountered by the learner that are known, or reasonably expected, to influence the course, nature, or extent of acquisition of some type of language "ability" on the learner's part. Within this very broad domain, Spolsky makes a major distinction between "formal" learning conditions, which he defines as "the various institutionalized

educational opportunities provided by a society for language learning" (p. 26); and "natural" or informal learning conditions, which are "available in different kinds and amount according to social conditions which determine the potential opportunities for a learner to interact with speakers and writers of the target language [in other than formal instructional settings]" (p. 26).

With regard to the measurement concerns at issue in attempting to operationalize process variables generally, it would appear useful to adopt the position that any and all language learning-related influences that impinge on the language learner, on an "around-the-clock" basis, and without regard to whether they are formal or informal in nature, would by definition fall into the process-variable category and should ideally be accounted for in a thoroughgoing description/analysis conducted for research purposes. The relative proportions of formal and informal conditions in given teaching/learning settings would be expected to vary appreciably across settings. For example, an American teenager taking an intermediate French course in a midwestern high school might be expected to encounter predominantly formal instruction in the language, with essentially no opportunities for informal learning exposure, except possibly for an occasional televised (and English-subtitled) French movie on a public television channel. In another contrasting setting, an American shipping agent in Marseilles would be continually exposed to a plethora of informal language exposure and use opportunities, both in the workplace and during available leisure hours; formal instruction, if any, might be restricted to the occasional self-paced study of correspondence course materials on grammar or vocabulary development. An "around-the-clock" perspective on describing, measuring, and analyzing the process variables operating on individuals or groups of language learners in each particular teaching/learning setting would help to keep researchers mindful of the numerous influences potentially operative in that setting in addition to any formal classroom instruction, as well as attentive to the need to take adequate account of these additional types of exposure in their research design, analysis, and results reporting.

With the significant exception of computer-assisted learning activities, which can be minutely tracked and logged with respect to both the stimuli presented and the nature of the learner's responses, it is, as a practical matter, impossible to fully specify the nature, quantity, and quality of the "process" variables that impinge on the learner in the course of even the most modest and circumscribed classroom learning activities, not to mention long-term intensive programs and informal learning situations. In any event, notwithstanding the complexities involved, it should be possible to determine and report with some degree of objectivity and rigor at least the most salient characteristics of given teaching programs or learning environments, insofar as these would be

expected to affect the observed learning results. One potentially workable approach would be for interested researchers to collaborate in the development and promulgation of a reasonably detailed but administratively feasible taxonomy of the major components of "language learning situations" generally, including language contact and use opportunities in both formal and nonformal settings and on a 24-hour basis. Subsequently, the agreed-upon taxonomy would be used as a means of objectively situating a given study with respect to its major "process" characteristics, as these are found to map onto the overall taxonomic framework.

Analysis Approach

The effective statistical analysis and interpretation of the input-process-output variables at issue in the Spolsky model and of their functional interrelationships would appear to involve two distinct but complementary approaches. The first of these, which might be termed the "integrative analysis" approach, would involve the use of appropriate analytic methods and procedures to integrate the data and findings of a number of separate research studies into the overall framework of the Spolsky model. Although no single research effort of workable scope could possibly encompass all of the variables at issue in the total model, a viable alternative would be to make use of analysis methodologies that are capable of consolidating the findings of a large number of individual research studies, each study focused on the relationships among those variables subsumed under a particular condition for learning. The goal of this approach would be to integrate available findings into rather robust estimates of the relationships between any two variables hypothesized on the basis of the Spolsky model.

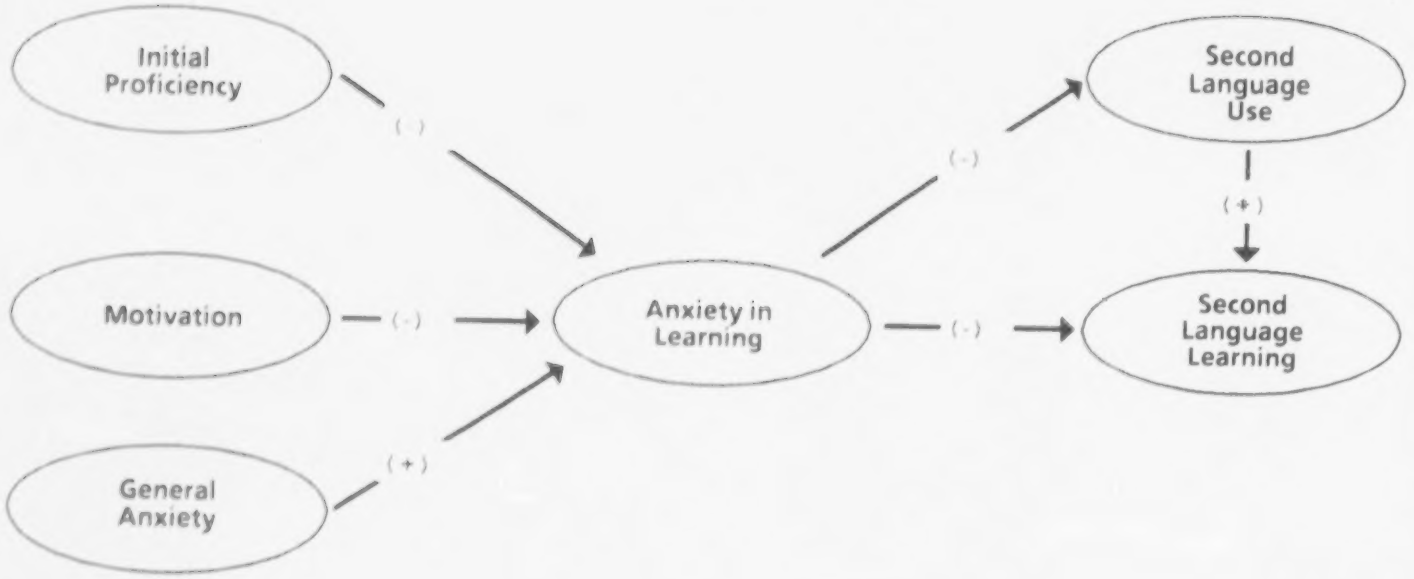
The second proposed analytic procedure, which is usefully characterized as the "model analysis" approach, would go beyond the bivariate analyses at issue in the integrative analysis approach, in an attempt to assess the higher-order patterns of interrelationships among many, if not all, of Spolsky's learning conditions. Consider, for example, his Condition 33:

> Condition 33. Second Language Learning Anxiety condition (typical, graded): Some learners, typically those with low initial proficiency, low motivation, and high general anxiety, develop levels of anxiety in learning and using a second language that interfere with the learning. [115]

While the necessarily general level of definition in Spolsky's description of this condition allows for alternative interpretations of the implied variable interactions, one likely set of interrelationships is displayed in Figure 2, which posits the linkage of six variables: positive relationships between General Anxiety and Anxiety in Learning, and between Second Language Use And Second Language Learning; and inverse relationships linking Initial Proficiency and Motivation on the one hand with Anxiety in Learning on the other; and between Anxiety in Learning and Second Language Use and Second Language Learning. The integrative analysis approach would be applied to estimate the nature and magnitude of the bivariate linkages among each of the individual variables considered to be subsumed under the general "Second Language Learning Anxiety" condition. However, such analyses would constitute only a very coarse overall test of the model implied in that condition for learning, and the model would not be supported unless all six hypothesized relationships were found both to exist and to be in the predicted direction. A more powerful test would require assessment of the overall fit of the predicted constellation of relationships with the obtained data. The program of analyses proposed in the model analysis approach would be directed at conducting such tests within and, to the extent possible, across the conditions for learning that specify causal dynamics.

A full explication of the issues and methods involved in conducting the two lines of integrative analysis and model analysis with respect to Spolsky's learning model is considerably beyond the scope of this paper. Moreover, the final definition of these issues and methods must await the resolution of basic design decisions regarding these efforts, as well as the further specification of the operational dynamics depicted in the Spolsky model. However, it is at this juncture possible to provide an overview of the proposed analysis program, with an emphasis on the associated organizational mechanisms that would need to be established and implemented to permit the efficient and effective conduct of these analyses.

FIGURE 2
Variable Dynamics Implied by Spolsky's Condition 33



Integrative Analysis Approach

The integrative analysis approach would involve carrying out a program of quantitative analyses of existing data and research findings, with these analyses aimed at subsequently supporting the model analysis approach by providing a closer operational specification of the models of language acquisition contained within Spolsky's learning conditions, more rigorous estimates of the model's parameters, and findings informing the further development of measurement and methodological guidelines for subsequent data collection and analysis. The analytic approach best suited to these purposes is meta-analysis.

Meta-analysis (Cochran, 1954; Glass, 1990; Light & Smith, 1971) involves the application of quantitative methods to synthesize the results of separate research studies in a particular area of inquiry. In its function, a meta-analysis parallels, at the statistical level, the verbal integration of the findings of various individual studies that is characteristically carried out in the typical "literature review" article. In the latter, the reported findings are qualitatively examined, with the goal of identifying lines or networks of consensus among them. Inconsistencies among the findings of individual studies are usually addressed through an attempt to relate these inconsistencies to differences in the makeup of the sample, the procedures followed, the assessment instruments administered, and so forth. In a meta-analysis, the same logical operations are performed, but through the use of quantitative rather than qualitative methods. For each effect examined across the studies comprising the meta-analysis, a statistical estimate is made of the value of that effect by pooling the results of all the studies in the analysis. Additionally, differences in the studies' estimates of the effect are statistically related to the characteristics of the studies to evaluate how those characteristics may have influenced the studies' results.

The statistical approaches used in meta-analysis are generally similar to those employed in single, "stand-alone" research studies, including analysis of variance and the various forms of correlation and regression analysis. However, while the unit of analysis for individual studies is typically the individual language learner, the unit of analysis for a meta-analysis conducted across several studies is the individual research studies themselves. Although not yet applied extensively to second language acquisition studies, meta-analysis has, over the past decade, seen extensive and highly successful use in other fields and disciplines. Within medical research, for example, meta-analysis has become the accepted procedure for integrating the results of clinical drug trials and for analyzing the efficacy of various forms of medical treatment (Mann, 1990). Within the education field, the meta-analytic

approach has been applied to such issues as the effects of class size (Glass & Smith, 1979), instructional approach (El-Nemr, 1979), and "open education" practices (Hedges, Giaconia, & Gage, 1981) on student achievement.

Even though most of the significant applications of the technique have occurred only within the last fifteen years, meta-analysis has matured into a broadly accepted and routinely applied procedure, due in large part to its foundation in standard quantitative and statistical methods and to its ability to meet the increasing need of the research community to assimilate and draw overall conclusions from the burgeoning number of individual research projects and project findings. The dramatic increase in the scope and diversity of second language learning research activities manifested over the past several years makes this field of inquiry an excellent candidate for meta-analytic attention.

The meta-analytic process can be viewed as proceeding through three distinct steps: compilation of existing results, statistical codification of compiled results, and analysis of the quantified data. Each of these steps is considered below with respect to the manner in which it might be implemented within a second language acquisition research context.

Compilation of Existing Results

As with any statistical analysis procedure, the ultimate value of a meta-analysis depends on the quality of the data upon which the analysis is performed and on the degree to which these data reflect the conditions to which one wishes to generalize the study results. Careful planning and execution in locating and compiling the results of all prior research pertinent to the question under investigation is critical to the success of the analysis.

The principal challenge in the compilation of prior research findings is to ensure that the assembled database is complete. It is not sufficient that the meta-analysis include the results of studies considered the most important or the "best" in the area under investigation, or that the analysis examine only those which are most directly amenable to quantitative analysis, or those published in the easily-accessible literature, or even those published at all. The application of any such criteria, either intentionally or inadvertently, in selecting findings for inclusion in the meta-analysis runs the risk of systematically influencing the types of data included and, by the same token, of seriously biasing the analysis results. Instead, the collection of research results should be as inclusive as possible, leaving to the capabilities of meta-analytic techniques the proper resolution of

questions as to how across-study differences in research techniques, quality of experimental control and measurement, the nature of publication, and so forth may have influenced the obtained results.

For example, the issue of published vs. unpublished research has been widely recognized as a problem in integrating and interpreting research findings, whether through meta-analysis or other means. To the extent that studies that obtain positive results (i.e., results supporting the existence of a relationship between two variables of theoretical interest) are more likely to be submitted and accepted for publication than those reporting negative results, the published literature--as well as any meta-analyses based on this literature--would overestimate the actual underlying relationships among the examined variables. Similar effects would be expected to stem from the reduced likelihood of publication of studies showing only weak relationships among variables, or of studies whose findings are at variance with previously published results. Although recently-developed procedures are available to estimate the effect of publication bias and to correct meta-analysis findings for that estimated effect (Hedges & Olkin, 1985; Rust, Lehmann, & Farley, 1990), the first, and essential, line of defense against publication bias is to make all possible efforts to locate unpublished as well as published research results through such means as announcements in professional publications, inquiries circulated through professional societies, formal or informal networks of interested researchers, and so forth.

Within the second language learning field, locating, obtaining, and using prior research results for meta-analytic or other purposes is considerably complicated by the interdisciplinary and international nature of this research, which makes it necessary for investigators to extend their inquiries across a much wider range of publications, professional associations, and other informational networks than would otherwise be the case. Further, the many different languages on which acquisition research is conducted (and in which the results are reported) impose considerable burdens on data gathering and increase the complexity of complilation and analysis.

Codification of Compiled Results

Two types of information about the located studies must be quantified before the meta-analysis can be conducted: the obtained results of the studies, and those characteristics of the studies thought likely to have influenced the study results or otherwise deemed to have potential value in serving the purposes of the meta-analysis.

The results obtained in each study used in the meta-analysis must be converted to a common metric that will allow their quantitative

integration. It is to be expected that the results of separate empirical efforts will have been reported in terms of a variety of inferential and descriptive statistics and with varying degrees of thoroughness. For example, in one study, the measured relationship between two variables may be reported only in terms of an obtained *F* value with its associated degrees of freedom and probability; a second study may present the same relationship solely through group means and standard deviations; while a third may provide this information in the form of a *t*-test between extreme groups, or as an obtained correlation coefficient; and so forth. For research studies that are essentially qualitative in nature, codification approaches other than simple statistical transformation are required. For example, for an investigation which provides its results only in the form of detailed narrative accounts of a number of case studies, the verbally-described research outcomes would need to be translated into suitable quantitative terms, such as the percentage of individuals reaching a particular performance criterion as deduced from the verbal descriptions. In any such instances of qualitative-to-quantitative conversion, steps must be taken to verify an adequate degree of coding reliability, through inter-judge consistency checks or other appropriate means.

Data Analysis

As previously indicated, the meta-analytic process is comprised of two separate analysis phases. The first seeks to determine the nature and magnitude of the relationship among variables that can be concluded from the results of the examined studies. In a relatively simple case, these findings may be manifested as, for example, the observed relationship between a dichotomous input or process variable (e.g., field dependent vs. field independent, trained vs. untrained, male vs. female) and a single continuous output variable (e.g., measured language proficiency). In such a circumstance, the usual measure of effect--the standardized means difference between the two groups--would be applied, using standard inferential techniques to test the null hypothesis of no difference between the two groups. With the rejection of the null hypothesis, the pooled measure of effect would stand as an estimate of the effect's size. Similarly straightforward are instances in which the study results consist of estimated correlations between two variables, with greater challenges being introduced when observed values on the independent variable are contrasted among multiple, qualitatively different conditions (e.g., alternative training approaches, different languages). Under such conditions, analysis of variance or least square approaches might be employed to test the effect of the distinct conditions in both their direct effects and, if required by the

situation, their interaction effects in combination with the other conditions or parameters.

Approaches used in the second phase of the meta-analysis--investigation of the relationship of particular characteristics of the individual studies to their reported results--must be equally diverse. In addition to considering how each characteristic relates to variance in the obtained findings, attention must be directed to the potentially confounding covariation of the context or framework within which the studies are conducted and to their obtained results. For example, it would be extremely useful to be able to combine and jointly interpret those research findings based on quantitative analyses of standardized data on relatively large samples of language learners and those derived from more naturalistic studies involving in-depth, qualitatively-oriented observation of small groups of individual learners. Confounding elements are, however, likely to intrude in attempting to do so. Large-scale quantitative studies are, by their very nature, more likely to have been carried out in those settings that facilitate the collection of such data, in particular, during formal language instruction in institutional settings. Such studies, as a result, may be expected to concentrate more on mechanisms and processes associated with formal language learning than on those involved in naturalistic language acquisition. Additionally, the population of language learners receiving large-scale instruction may be presumed to differ in a number of important ways (for example, age, attitude toward and degree of motivation for language learning) from individuals who acquire second language skills primarily through non-institutional experiences. As a consequence, differences in the particular research methods employed may be expected to be confounded with differences in the instructional dynamics involved in the two learning contexts, as well as with the particular learner populations involved. Considerable analytic attention and effort, which may require bringing to bear the full gamut of scaling and statistical analysis procedures currently available for the purpose, will be needed to account and adjust for these and other potential confounding effects in conducting the proposed meta-analyses.

It should also be acknowledged that pragmatic limitations associated with the language teaching/learning endeavor itself (at least in its present-day forms and contexts) may be expected to reduce the interpretive scope and applicability of even the most extensive and carefully conducted meta-analyses. First, most of the reported research on language learning is conducted within fairly restrictive institutional frameworks (e.g., typical school-based language programs), with the consequence that the number and nature of the independent variables investigated--as well as the range of variation manifested in these variables--are for all practical purposes limited to those which occur normally within these particular settings or can be reasonably

engineered by the researcher within the physical and administrative limitations of these settings. In addition, in most real-life learning situations, language instruction is not imposed as an experimental treatment upon a truly random sample of learners; as a consequence, the learning studies perfomed in these settings must be considered observational in nature rather than experimental in the strict psychometric sense of the term. Effective randomization-based control on the sample characteristics of learners involved in informal, non-institutionalized types of language learning is even more problematical, given the overwhelmingly self-selected nature of this population and the consequent virtual certainty that those who acquire competence in one or more non-native languages through informal learning processes differ in significant ways from the general population of potential language learners.

As a consequence of these and other considerations, it is extremely important to recognize that, however broad the research base used in a meta-analysis may be, there will always be a finite ability to generalize its conclusions to the overall theoretical population. The crucial question of how, in specific, quantifiable terms, individuals who undertake to learn languages differ from the universe of potential language learners is an issue that should be addressed as a priority matter both in reviewing already-conducted studies and in carrying out research intentionally targeted at the further elucidation of this question. By the same token, a variety of other non-random factors operative in the current body of language learning and acquisition research studies (including under- or over-representation of particular languages or language families; the predominance of particular models of or approaches to language instruction to the exclusion of other possible models/approaches; and so forth) should be addressed, with the dual purpose of appropriately qualifying the meta-analytic conclusions drawn from available studies and of identifying desirable directions for future research.

Application to the Spolsky Model

As is no doubt clear from the preceding discussion, a program of meta-analyses spanning the broad scope of Spolsky's learning model would require not only extensive effort on the part of a large number of qualified researchers but also extensive coordination of that effort. Such coordination would in all probability need to be accomplished, on an ongoing basis, under the auspices of a research institution or professional organization with established ties throughout the language learning research community (both in the United States and abroad),

and with an administrative infrastructure adequate to the tasks involved.

The functions that such a coordinating organization would need to undertake extend throughout the steps of the meta-analysis procedure. With respect to the *overall planning and management* of such an undertaking, important organizational functions would include defining the overall objectives, scope, and schedule for the analysis program; enlisting the support of other language learning researchers and research organizations; and establishing the vehicles by which the program's efforts would be consolidated and its interim results communicated to the language teaching and research community.

To facilitate the *assembly of research results* for the meta-analysis efforts, the organization would probably need to establish and operate a clearinghouse or registry of existing research reports and articles. In addition to soliciting, receiving, and managing the collection of such publications, the organization would undertake to develop and publish standards for the acceptance of reports and articles into the registry (e.g., degree of relevance to further investigating or explicating Spolsky's conditions for learning model, adequacy of experimental design, level of empirical content, and so forth). More particularly, the coordinating organization would have the important responsibility of coding and categorizing the reported findings with respect to the particular Spolsky condition(s) to which they are considered to relate. This function would require the prior establishment of relevant coding systems and procedures, including development of the closest possible operational definitions of these conditions, desirably through the use of a Delphi process involving the participation of a large number of interested and qualified researchers.

The coordinating organization would support the *coding of assembled study results* for entry into suitable meta-analyses through:

(a) establishing and promulgating standard measurement procedures to be used to combine research results in the meta-analyses;

(b) establishing guidelines for reporting research results that will in the longer term facilitate their meaningful aggregation via the meta-analytic process;

(c) identifying which study characteristics it is most important to examine across the meta-analyses;

(d) developing coding schemes and associated definitions and procedures for quantifying study characteristics;

(e) supporting the conduct of interim analyses to uncover any areas of unreliability developing in coders' ratings;

(f) providing the mechanisms for modifying the coding techniques to correct any uncovered deficiencies.

There are also at least two functions that the coordinating organization could perform to assist the *data analysis efforts* within the meta-analyses:

(a) establishing a master list of study characteristics to be considered for possible analysis in each meta-analysis;

(b) putting into place procedures and vehicles to facilitate coordination in the conduct of related meta-analyses, to ensure complementarity of procedures and the mutual utility of results.

Finally, the coordinating organization should play a central role in *integrating and communicating the meta-analysis results* as they are being developed and ultimately formally disseminated. Possible mechanisms include the sponsorship of periodic meetings and workshops for participating researchers, the maintenance of an electronic bulletin board or other communications network, and technical and operational assistance in publishing and distributing the obtained results.

Model Analysis Approach

The substantive meta-analytic results of the integrative analysis approach as described above will directly benefit the conduct of analyses aimed at empirically testing causal hypotheses derived from the various facets of the Spolsky model itself. First, the meta-analysis results should allow the required scope of inquiry to be at least somewhat narrowed. For example, if convincing evidence emerges from the meta-analyses that a particular variable currently at issue within the Spolsky conditions plays no appreciable role in language learning and acquisition, or that two or more variables co-vary so closely and consistently that each adds no useful individual variance in the presence of the other(s), it should be possible to correspondingly reduce the total number of variables requiring consideration in the model analysis approach. Given the very considerable data requirements of the model analysis approach in general, any possible reduction in the number of variables and measures involved will result in substantial savings in overall research time and effort, as well as reduce analytic and interpretive complexity.

Results of the meta-analyses will also provide substantial guidance in further specifying the procedures to be used to test various components of the Spolsky model. As discussed in the following section, a critical step in preparing to test models of language learning based on Spolsky's inventory of conditions is to develop detailed specifications of the presumed variable relationships at issue within these conditions. The general structure described in a given condition for learning must be refined into more specific sets of hypothesized, empirically measurable relationships before a fit can be attempted to actual data; this effort can be considerably aided by carefully considering available meta-analytic results, in conjunction with theoretical and logical analysis.

The pattern of obtained meta-analysis results will also help to prioritize the factors to be considered in model testing with respect to particular language tasks, settings, and conditions. Because of the wide-ranging scope of the Spolsky model, the number of possible factors to be considered is extremely large, and given the practical limitations associated with any research endeavor, it will without doubt be necessary to focus even the most ambitious studies more narrowly than the total Spolsky model would require. Results of the meta-analyses could provide very useful input to guide those decisions.

The statistical analytic method that would at this point appear to have the greatest potential for properly and effectively testing alternative causal models based on Spolsky's conditions is *covariance structure modeling*. Based on the analytic procedures established in the work of Jöreskog and van Thillo (1972), Keesling (1972), and Wiley (1973), this family of techniques, commonly referred to as the JKW model, is contained in, and most readily available through, the LISREL analysis package. Because of this close association, we shall use the term "LISREL modeling" to refer generically to these techniques, which collectively show considerable promise for handling the required analysis tasks, even though other still more generalized statistical approaches, such as McArdle's (1978) reticular action model, may ultimately prove more useful.

In its operation, LISREL modeling consists of two components, each of which has associated requirements and implications for the analytic enterprise involved in testing the Spolsky model. The first of these, referred to as the Measurement Model, describes the relationship between the observed measures in a given analysis and the underlying latent variables (corresponding to the theoretical concepts under investigation) that the observed measures are consided to reflect, albeit at least somewhat imperfectly and indirectly. The second component, known as the Structural Equation Model, is subsequently applied to test the hypothesized relationships among the latent variables.

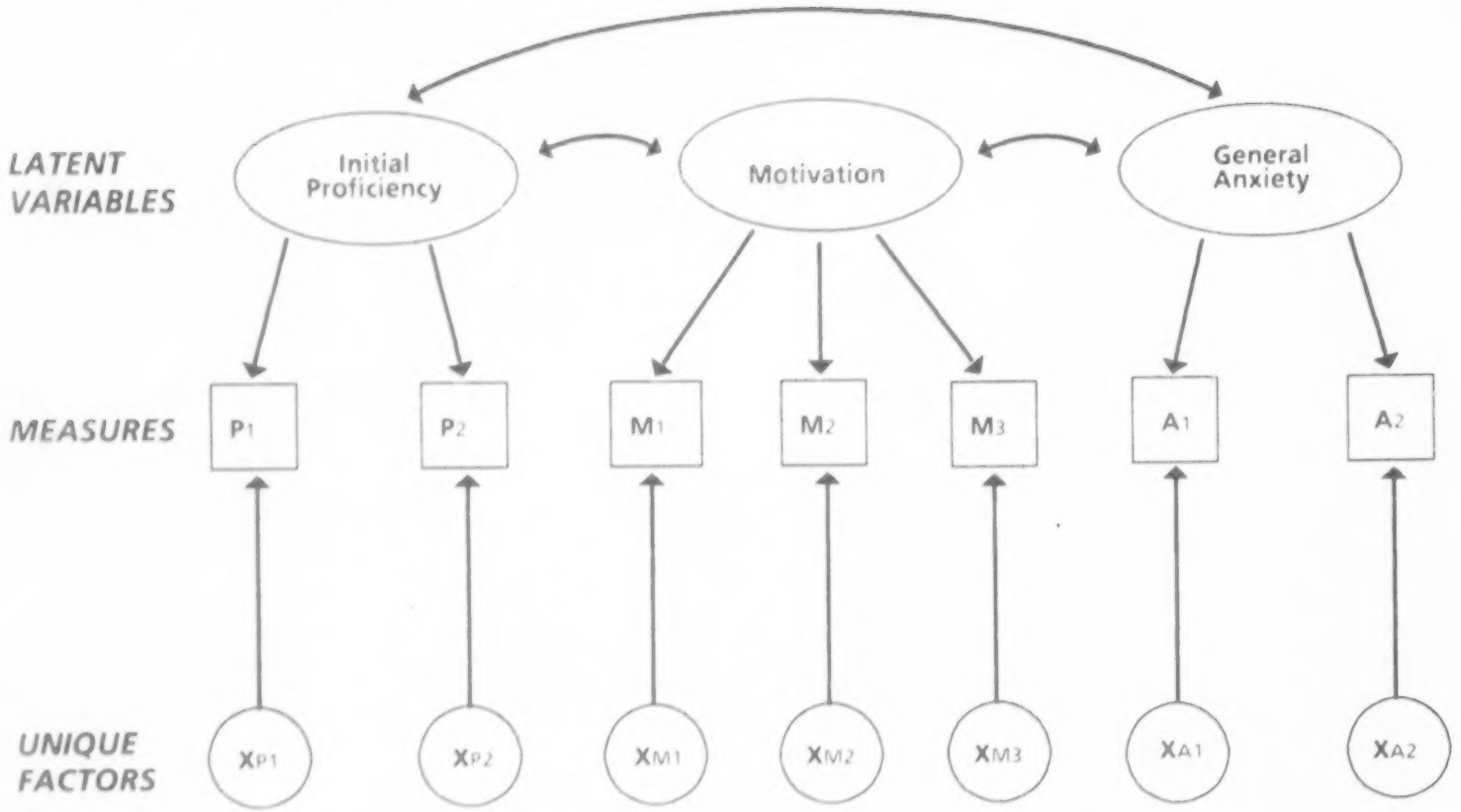
The Measurement Model

Theoretical concepts, and their associated latent variables, can of course never be directly observed in research; their individual effects, as well as their interrelationships with other variables, can only be inferred from scores on the various tests, questionnaires, or other types of measures (all known as "indicators" in LISREL terminology) that the researcher incorporates in the study to embody and reflect the latent variables of interest. To take an example, Figure 3 depicts the three latent variables--Initial Proficiency, Motivation, and General Anxiety--that are characterized by Spolsky as potentially influencing Second Language Learning Anxiety (Condition 33; p. 115). In a typical research design, each of these postulated latent variables would be represented by one or more measurement instruments, shown as square boxes in Figure 3. The measures nominally associated with a given latent variable can be expected to be statistically correlated with other measures associated with the latent variable. Each measure's unique variance, on the other hand, reflects the influence of other, unmeasured variables (assuming no covariation among sources of measurement "error"). Covariation among measures in different sets (not indicated in the figure for simplicity) would reflect the association between the corresponding latent variables or similar dynamics.

The quantitatively oriented reader will recognize the structure shown in Figure 3 as a factor analytic model; indeed, the analysis techniques employed to establish a measurement model using the LISREL approach are based upon confirmatory factor analysis (CFA). By applying CFA techniques to the data derived from measures ostensibly reflecting the latent variables of interest, the researcher can test the fit of different hypothesized measurement models to the observed patterns of covariation in the obtained data. For example, the researcher may decide to constrain the model to reflect known differences in the quality of the measures, or to test the implications of known common sources of measurement error (e.g., methods variance).

Clearly, to profitably conduct such analyses on the wide range of measures needed to fully examine the Spolsky model would require considerable standardization of the measures used in this effort, as has already been emphasized. Using a multitude of study-specific measures to assess given individual model variables would constitute a grave, or perhaps even fatal, obstacle to the proper empirical analysis of the Spolsky (or for that matter, any other) learning model. This would be especially true if--as is likely--the use of particular measures were limited to particular research settings, particular learner populations,

Figure 3
The Measurement Model



or the work of particular researchers. Such a disparity of measures employed across a number of individual studies would make meaningful assessment of the covariation among the alternative measures a virtually impossible task.

A formal agreement among language researchers to employ identical tests or other standardized protocols across a number of different research contexts and learner populations would be a laudable and necessary first step toward uniform and interoperable data gathering and analysis, but would not in and of itself address the numerous validity-related issues involved in operationally defining the variables themselves. Consider, for example, such an apparently straightforward variable as "level of education." On first thought, this would seem to be easily and properly expressed as the number of elementary/secondary grade levels successfully completed or, at higher levels, postsecondary degrees received. However, closer consideration of the underlying trait this particular measure is intended to reflect may raise questions about the validity of such a scale. If "level of education" is intended to reflect years of acclimation to an academic setting, then a straightforward grade-level scale might be considered an acceptable and adequate metric (although correction for any repeated years of formal schooling would be necessary for complete accuracy). If, on the other hand, this variable were meant to reflect some communally achieved body of knowledge on the part of the study participants, then geographical location, socioeconomic status, school quality, and a host of other sub-variables not accounted for in the conventional grade-level scale but potentially associated with acquisition or non-acquisition of this body of knowledge would be inadequately represented.

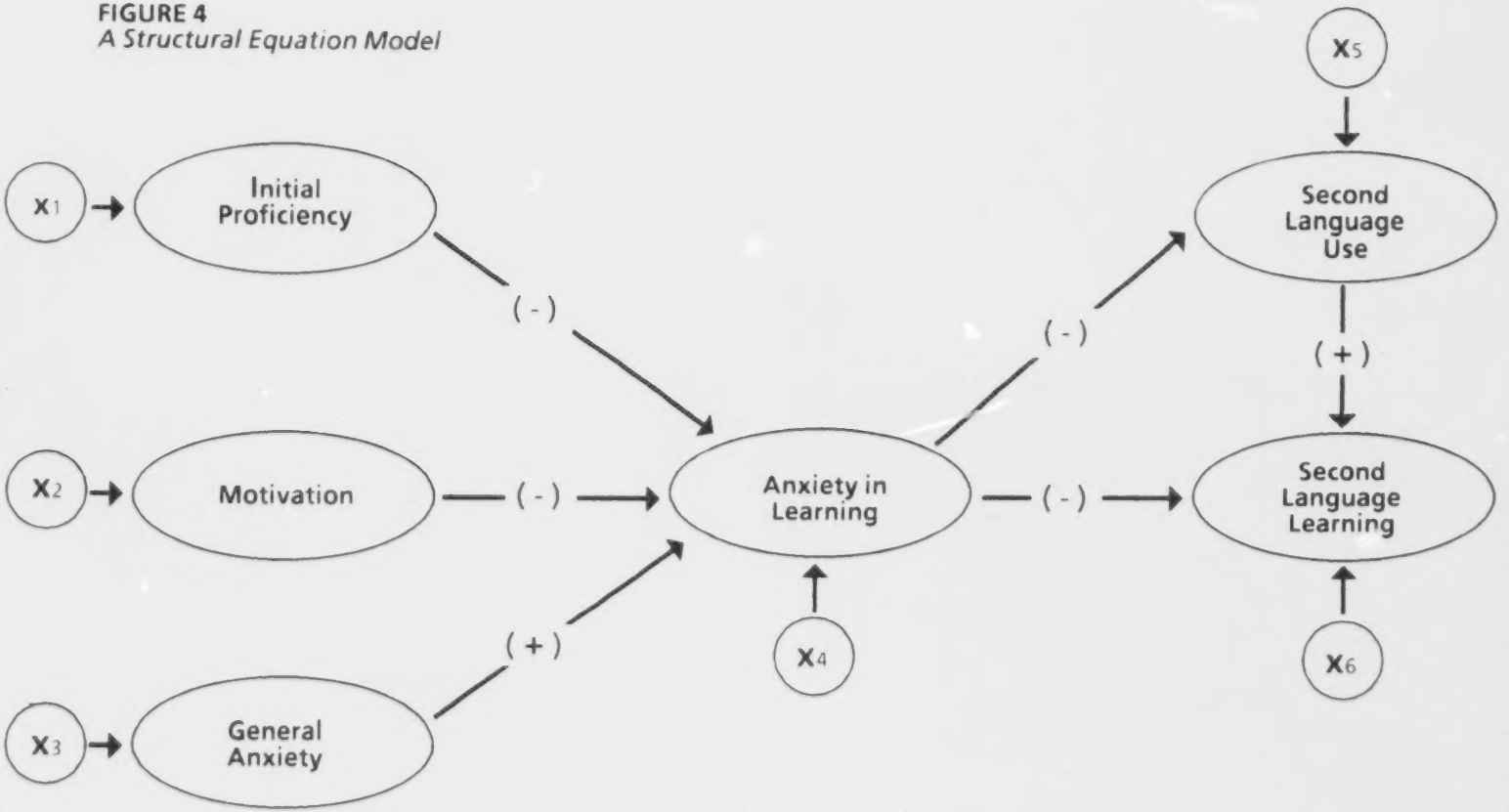
What the above example suggests (and which would be even more apparent in the case of still more complex variables) is that a considerable amount of intensive, coordinated activity on the part of a large number of foreign/second language researchers will be needed in order to propose, and obtain an adequate consensus on, both the detailed operational definitions of the major variables at issue in the Spolsky model and the most suitable instruments and measurement approaches to quantifying the variables so defined. Joint communication and combined effort across research sub-fields that have not traditionally been in particularly close collaboration (first and second language learning; foreign language and English as a second language learning; formal learning and informal acquisition; experimental and ethnographic research) will be necessary if any significant headway is to be made. Only through this degree of coordination and collective involvement will it be possible to define the parameters of, and develop adequately valid and sufficiently robust measures for, the numerous input, process, and output variables that must be considered in conducting effective research studies across the many present and

prospective teaching/learning contexts in which language learning does (or could potentially) take place.

The Structural Equation Model

While the measurement model just discussed performs the useful, and indeed necessary, first task of bridging the gap between the recorded data and the theoretical constructs hypothesized to underly them, the structural equation model addresses the further critical step of empirically testing the hypothesized relationships among these constructs. To accomplish this, one of several alternative techniques (whose applicability varies according to the nature of the data and of the hypotheses being tested) is applied to assess the goodness of fit between the hypothesized variable relationships and the observed pattern of covariation among the measures representing the variables. To exemplify and explain this approach, let us continue the example of a set of data hypothetically collected to test Spolsky's Condition 33, in which each theoretical variable has been perfectly and completely assessed by a single measure. In this ideal situation, the structural model for Condition 33 (as described earlier) would replicate Figure 2 with only the addition of a "model error" term for each of the six variables (see Figure 4). Each of the error terms represents the total influence on the corresponding variable of all variables that in reality affect that variable but are not included in the tested model (commonly referred to as exogenous variables). As in the measurement model, different assumptions about the nature and degree of relationship among the exogenous variables can be tested as part of fitting the model to the data. Within this broad analytic framework, a variety of alternative hypotheses can be tested. For example, one can statistically constrain particular parameters of the model to be equivalent or to assume particular values in order to assess how such manipulations affect LISREL estimates of unconstrained model parameters and of the overall fit of the model. Additionally, proper manipulation of LISREL procedures makes possible the inclusion of hypothesized interaction effects, nonlinear relationships, model differences across groups of subjects, and the impact of variables across time.

FIGURE 4
A Structural Equation Model



Application to the Spolsky Model

Clearly, the power of LISREL techniques to simultaneously consider complex relationships among a number of variables, and the flexibility of these techniques in modeling many different forms and patterns of variable relationships, make the LISREL approach ideally suited to the analytic processes at issue in evaluating the Spolsky learning model. However, before large-scale, LISREL-based analyses of the Spolsky model can be conducted, much coordinated preparatory work will have to be accomplished. Specifically, in addition to the previously-discussed development and implementation of standardized measurement procedures and the accumulation of a large body of data based on these procedures, further development will be required on the theoretical front before meaningful and useful LISREL analyses can be undertaken.

In its total, "macro" form, the Spolsky learning model is not amenable to adquate quantitative analysis as a unified conceptual entity, since it is simply too complex and extensive to permit any single test of all of the hypothesized relationships that can be derived from it. The data collection requirements of such an analysis are well beyond the limits of practicality, and even if such an extensive data base were available, the associated statistical analysis requirements would dwarf the capacities of any currently available analysis programs or procedures. At least for the foreseeable future, it would appear that any operationally viable quantitative research program aimed at further analyzing/validating the Spolsky model will need to adopt the strategy of initially conducting a number of smaller-scope research studies, each addressing a particular componential sub-model of the overall Spolsky model, leaving to subsequent analysis phases the question of how the sub-models interrelate within the larger "macro" framework.

One obvious approach for dividing the Spolsky model into componential models would be to take the individual learning conditions as the conceptual structures for defining the particular sub-models. Setting aside for this purpose those learning conditions that are definitional in nature (e.g., Condition 68, "Classroom condition [typical of formal learning, graded]: The learning takes place in a closed physical space, a single classroom" [p.24]), each of Spolsky's conditions for learning can be viewed as a model of a particular causal dynamic hypothesized to contribute to or detract from language learning and/or acquisition. Although this would be a fairly reasonable overall approach, caution would need to be exercised to guard against damaging the power of the Spolsky model through an overly rigid segmentation of the model by learning condition category. It may therefore be desirable,

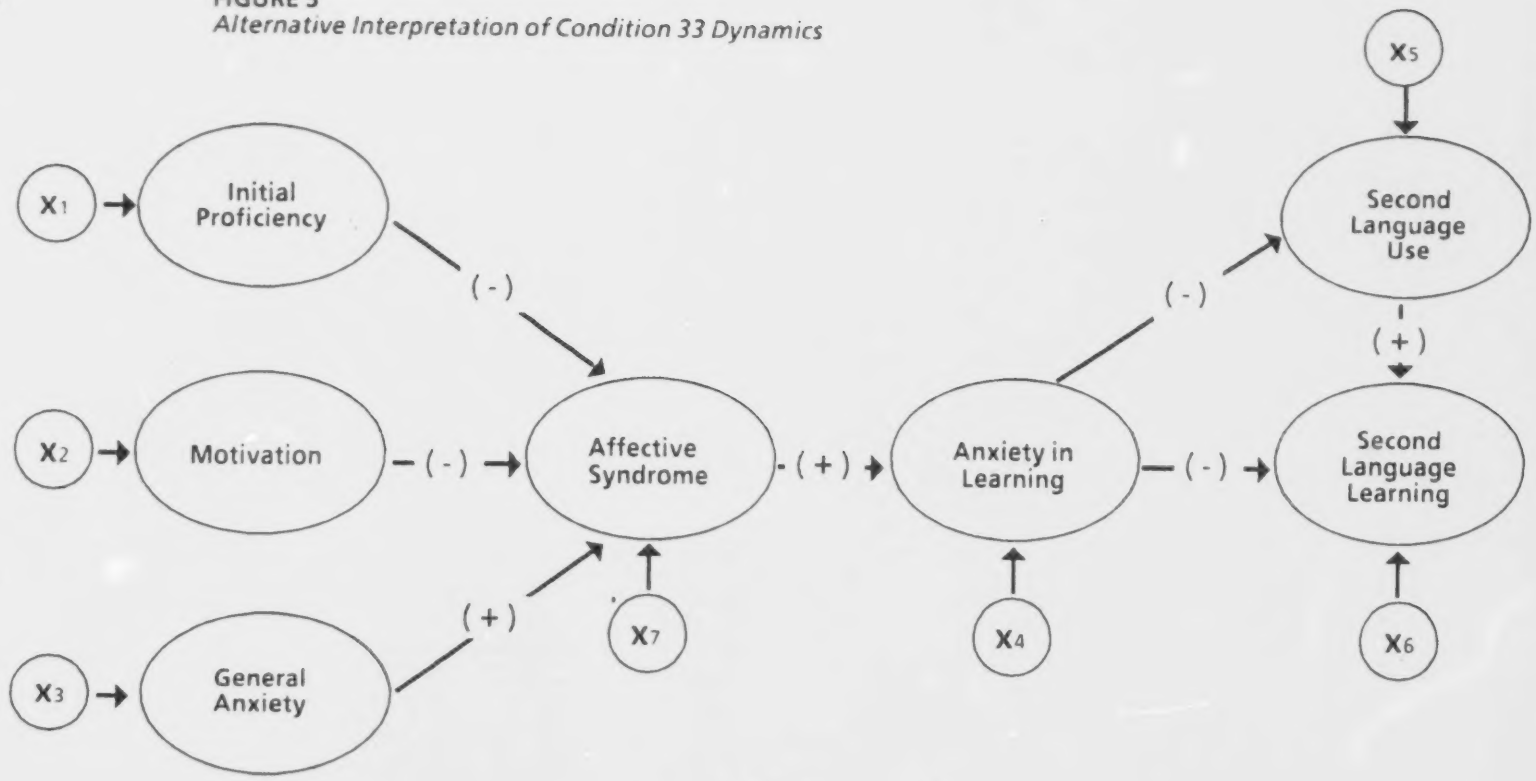
in at least some cases, to define sub-models that allow for the investigation of language learning processes across highly-related conditions, thereby preserving the integrating dynamics thought to influence language learning processes across as well as within the individual Spolsky conditions. Determining the specific composition of the sub-models to be examined through the LISREL procedures, as well as the chronological sequence in which they would most advisedly be addressed, would constitute a further, and very substantial, role for the proposed coordinating organization previously discussed.

A second major effort that will need to precede effective LISREL analysis of the Spolsky model is the further specification, in precise operational terms, of the latent variables at issue and their posited functional interrelationships. For example, the portrayal in Figure 2 of the variable dynamics described in Condition 33--showing Initial Proficiency, Motivation, and General Anxiety as independently influencing Anxiety in Learning, which in turn influences Second Language Use and Second Language Learning--is only one of many possible interpretations of the processes at issue in Condition 33. An alternative and equally defensible interpretation of Spolsky's definition of this condition is shown in Figure 5, in which the variables of Initial Proficiency, Motivation, and General Anxiety jointly form a syndrome that in turn influences Anxiety in Learning.

The difference between the models shown in Figures 2 and 5, though seemingly small, is theoretically and analytically quite substantial. In the Figure 2 model, the three variables are considered to operate independently to influence Anxiety in Learning. If the relationships implied in this model were indeed true, then a measured decrease in Motivation would result in increased Anxiety in Learning, even in the presence of high Initial Proficiency and low General Anxiety. In the Figure 5 model, this would not necessarily be the case.

In addition to the structural arrangement of the described variable dynamics, efforts to more explicitly specify the Spolsky conditions of learning would need to consider how complete each sub-model is in describing the effects of the variables it subsumes. Again examining Condition 33, it can be seen that, considered by itself, the condition suggests that the effects of Initial Proficiency, Motivation, and General Anxiety on Second Language Learning operate solely through their direct (Figure 2) or indirect (Figure 5) effect on Anxiety in Learning. However, it is reasonable to presume that in addition to their effect on Anxiety in Learning, high Initial Proficiency and Motivation would directly promote Second Language Use. If, in fact, Initial Proficiency and Motivation do have a direct effect on Second Language Use, a model of Condition 33 that omitted those direct effects would overestimate the strength of their effect on Anxiety in Learning and the effect of that variable in turn upon Second Language Use.

FIGURE 5
Alternative Interpretation of Condition 33 Dynamics



A coordinated program of research on Spolsky's model of second language learning would benefit greatly from a common understanding of the precise variable relationships to be explored in relation to the conditions for learning, these being operationally specified at the level of detail required for proper measure selection or development and for data collection and entry into the LISREL analysis program. This is not meant to imply that any single interpretation of Spolsky's learning conditions must operate as the single impetus and guide for model testing and refinement. Rather, an effort should be undertaken to establish a fuller specification (or sets of alternative specifications) of the Spolsky conditions as a point of departure for further theoretical development and empirical research. Viable mechanisms for productive discourse and collective decisions on these issues throughout the language learning research community need to be established to support both the development of the model specifications and the analysis and dissemination of research results addressing the various sub-components of the Spolsky model. The profession can, and should, rise to the conceptual, analytical, and organizational challenges that Bernard Spolsky has so eloquently and so effectively posed on its behalf.

References

Bachman, L. F., & Savignon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal*, 70, 380-390.

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test.* New York: Psychological Corporation.

Cartier, F. A. (1980). Alternative methods of oral proficiency assessment. In J. R. Frith (Ed.), *Measuring spoken language proficiency* (pp. 7-14). Washington, DC: Georgetown University Press.

Clark, J. L. D. (1979). Direct vs. semi-direct tests of speaking ability. In E.J. Brière & F. B. Hinofotis (Eds.), *Concepts in language testing: Some recent studies* (pp. 35-49). Washington, DC: Teachers of English to Speakers of Other Languages.

Clark, J. L. D. (1981). Survey measures: Language. In T. S. Barrows (Ed.), *College students' knowledge and beliefs: A survey of global understanding* (pp. 25-35). New Rochelle, NY: Change Magazine Press.

Clark, J. L. D. (1988a). Toward a research and development strategy for computer-assisted language learning. *CALICO Journal*, 5, 5-23.

Clark, J. L. D. (1988b). Validation of a tape-mediated ACTFL/ILR-scale based test of Chinese speaking proficiency. *Language Testing*, 5(2), 187-205.

Clark, J. L. D. (1989). Multipurpose language tests: Is a conceptual and operational synthesis possible? In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics. Language teaching, testing and technology: Lessons from the past with a view toward the future* (pp. 206-215). Washington, DC: Georgetown University Press.

Clark, J. L. D., & Clifford, R. T. (1987). The FSI/ILR/ACTFL proficiency scales and testing techniques: Development, current status, and needed research. In A. Valdman (Ed.), *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency* (Committee for Research and Development in Language Instruction) (pp. 1-18). Bloomington: Indiana University.

Clark, J. L. D., & Lett, J. A., Jr. (1988). A research agenda. In P. Lowe Jr., & C. W. Stansfield (Eds.), *Language in education: Theory and practice. Second language proficiency assessment: Current issues* (Vol. 70, pp. 53-82). Englewood Cliffs, NJ: Prentice Hall Regents.

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrica, 10*, 101-129.

Ehrman, M. (1990). The role of personality type in adult language learning: An ongoing investigation. In T. S. Parry & C. W. Stansfield (Eds.), *Language in education: Theory and practice: 74.*

Language aptitude reconsidered (pp. 126-178). Englewood Cliffs, NJ: Prentice Hall Regents.

El-Nemr, M. A. (1979). *Meta-analysis of the outcomes of teaching biology as enquiry*. Unpublished doctoral dissertation, University of Colorado.

Gardner, R. C., Lalonde, R. N., Moorcroft, R., & Evers, F. T. (1987). Second language attrition: The role of motivation and use. *Journal of Language and Social Psychology, 6*(1), 29-47.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House.

Glass, G. V. (1990). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.

Glass, G. V., & Smith, M. L. (1979). Meta-analysis of research on the relationship of class-size and achievement. *Evaluation and Policy Analysis, 1*, 2-16.

Hedges, L. V., Giaconia, R. M., & Gage, N. L. (1981). *The empirical evidence on the effectiveness of open education* (Vol. II, Final report of the Stanford University Research Synthesis Project). Stanford, CA: Stanford University.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hilton, T. L., Grandy, J., Kline, R. G., & Liskin-Gasparro, J. E. (1985). *The oral language proficiency of teachers in the United States in the 1980's: An empirical study*. Final report for U.S. Department of Education Grant No. G008201398 (Phase I) and G008302366 (Phase II), "Self-Assessment of Oral Language Proficiency." Princeton, NJ: Educational Testing Service.

Jöreskog, K. G., & van Thillo, M. (1972). *A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables*. Princeton, NJ: Educational Testing Service.

Keesling, J. W. (1972). *Maximum likelihood approaches to causal flow analysis*. Unpublished doctoral dissertation, University of Chicago.

Lado, R. (1978). Scope and limitations of interview-based language testing: Are we asking too much of the interview? In J. L. D. Clark (Ed.), *Direct testing of speaking proficiency: Theory and application* (pp. 113-128). Princeton, NJ: Educational Testing Service.

Lambert, R. D., & Stansfield, C. W. (1988). *Development of a procedural model and associated testing instruments for the diagnostic measurement of language proficiency growth, attrition, and maintenance* (Final project report submitted to National Security Agency). Washington, DC: Center for Applied Linguistics.

Lantolf, J. P., & Frawley, W. (1985). Oral proficiency testing: A critical

analysis. *Modern Language Journal, 69*, 337-345.

Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review, 41*, 429-471.

Lowe, P. L., Jr., (1986). Proficiency: Panacea, framework, process? A reply to Kramsch, Schulz, and, particularly, to Bachman and Savignon. *Modern Language Journal, 70*, 391-397.

Mann, C. (1990). Meta-analysis in the breech. *Science, 249*, 476-480.

McArdle, J. J. (1978). *A structural view of structural models.* (Paper presented at the Winter Workshop on Latent Structure Models Applied to Developmental Data, University of Denver, Department of Psychology.).

Munby, J. (1978). *Communicative syllabus design.* Cambridge: Cambridge University Press.

Oxford, R., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal, 73*, 291-300.

Oxford, R. L. (1990). Styles, strategies, and aptitude: connections for language learning. In T. S. Parry & C. W. Stansfield (Eds.), *Language in education: Theory and practice: 74. Language aptitude reconsidered* (pp. 67-125). Englewood Cliffs, NJ: Prentice Hall Regents.

Pimsleur, P. (1966). *Language Aptitude Battery.* New York: Harcourt Brace.

Rust, R. T., Lehmann, D. R., & Farley, J. U. (1990). Estimating publication bias in meta-analysis. *Journal of Marketing Research, 27*, 220-226.

Spolsky, B. (1989). *Conditions for second language learning.* Oxford: Oxford University Press.

Stansfield, C. W., Kenyon, D. M., Paiva, R., Doyle, F., Ulsh, I., & Cowles, M. A. (1990). The development and validation of the Portuguese Speaking Test. *Hispania, 73*(3), 641-649.

Wiley, D. E. (1973). The identification problem for structure equation models with unmeasured variables. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences.* New York, NY: Seminar.

Authors

JOHN L. D. CLARK, Dean of Program Evaluation, Research, and Testing, Defense Language Institute, Presidio of Monterey, CA 93944-5006. Specializations: language testing, applied research, and technology applications in second language learning.

FRANCIS E. O'MARA, Technical Director, PRC, Inc., 12001 Sunrise Valley Drive, Reston, VA 22091. Specializations: second language acquisition research and program evaluation.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editors encourage the submission of manu-scripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to Editor, *Applied Language Learning*, Office of the PRovost, Defense Language Institute Foreign Language Center, Presidio of Monterey, CA 93944-5006.

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on $8\frac{1}{2}$ x 11 inch paper, double-spaced, with ample margins. Where feasible, manuscripts are preferred on disc (ASCII file or XEROX Viewpoint). Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and if necessary notes.

Author Identification The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating:

full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

PIN: 055744 $\frac{1}{N}$ 000
GPO 583-000/11080

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

205

Applied Language Learning

VOLUME 2 • NUMBER 2
1991

206

Applied Language Learning

Volume 2 • Number 2

1991

Defense Language Institute
Foreign Language Center

207

APPLIED LANGUAGE LEARNING

PB 65 - 91 - 2

1991 • Volume 2 • Number 2

Applied Language Learning (ISSN 1041-679X) is published semi-annually by the Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. *ALL* presents critical integrative articles and research reports on adult language learning for functional purposes.

COL Donald C. Fischer, Jr., Commandant
Defense Language Institute, Foreign Language Center
Department of the Army

This publication presents professional information: the views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material to meet space constraints.

This publication approved for public release: distribution is unlimited.

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the Institute. Cover design and graphics by Barney Inada, Visual Production, DLI; copyediting and formatting, Jack Boyd, Publications, DLI.

APPLIED LANGUAGE LEARNING

PB 65 - 91 - 2

1991 • Volume 2 • Number 2

EDITORIAL BOARD

Lidia Woytak, Editor	Curriculum Specialist , DLI
Kathleen M. Bailey	Head, TESOL/ESL Program, Monterey Institute of International Studies
David P. Benseler	Editor, *The Modern Language Journal*
John L. D. Clark	Dean, Program Evaluation, Research, and Testing, DLI
Ray T. Clifford	Provost, DLI
Annemarie Frazier, Word Processing Consultant	Academic Coordinator, School of Romance Languages, DLI
Martha Herzog	Dean, Curriculum and Instruction, DLI
Betty Lou Leaver	Dean, Slavic School, DLI
John A. Lett, Jr.	Director, Research and Evaluation, DLI
June K. Phillips	Director, Tennessee Foreign Language Institute
Gary D. Walter	Librarian, DLI

Forget About Needles in Haystacks!

Discover the information you need quickly and easily by consulting **Second Language Instruction / Acquisition Abstracts (SLIA)** — your window on the entire spectrum of second-language instruction literature published worldwide.

As a specialist in your field, it is important to keep pace with the latest research findings and teaching methods. **SLIA** can help you do just that with timely coverage of articles culled from among 1,900 serials, selected books, and enhanced bibliographic citations for relevant dissertations.

SLIA appears biannually — in July and December. Each issue offers approximately 700 abstracts, divided into 13 substantive areas, plus separate source, author, and subject indexes.

Among the subject areas covered are:

- Applied linguistics
- Non-native language pedagogy
- Non-native language acquisition
- Adult language development
- Bilingual education
- Computer-assisted learning, and more.

Second Language Instruction / Acquisition Abstracts

(a Division of Linguistics & Language Behavior Abstracts)

P.O. Box 22206 • San Diego, CA 92192-0206
(619) 695-8803 • FAX (619) 695-0416
SOCIO@SDSC.BITNET

APPLIED LANGUAGE LEARNING

PB 65 - 91 -2

1991 • Volume 2 • Number 2

CONTENTS

vii From the Editor

1 Language Pedagogy and Effective Technology Use
Nina Garrett

15 Training in Instructional Technologies: Skills and Methods
Sue K. Otto

31 Effective Strategies for Implementing Language Training Technologies
Robert Ariew

45 Managing Innovation and Change for Instructional Technology
Donald C. Holznagel

61 Software for Language Training: Directions and Opportunities
James C. Pusack

77 Hardware for Language Training: Coping with Confusion
Michael D. Bush

92 Information for Contributors

From the Editor

This issue of *Applied Language Learning* is devoted to the use of electronic technology in foreign language instruction, and is based on papers submitted to an invitational conference on "Improving Foreign Language Teaching through Technology," held at the Defense Language Institute on October 29-30, 1990.

The conference was held under the auspices of the Educational Technology Needs Assessment (ETNA), a research project undertaken at DLIFLC to investigate the principled and effective use of technology in language instruction. The invitational conference was convened by DLIFLC and the Institute for Simulation and Training, under an agreement with the Defense Training and Performance Data Center, which assisted DLIFLC throughout the ETNA project. The conference brought together a wide spectrum of university, government, and military participants. Certain ideas presented there have already been implemented in the DLI curriculum. Speakers were selected for the nature and extent of their contributions to the state of the art in foreign language education technology. Each presentation was followed by comments provided by two discussants. Following the conference, the presenters were invited to review and revise their papers in light of the discussants' comments.

Some foreign language educators still have mixed feelings about the use of computers in their programs, perhaps wondering whether computers are truly worth the effort and the expense. And there may still be a number of classroom language teachers who, passing a shiny new computer, cannot help wondering, "Are you eventually going to take my job away?" Ray Clifford, DLI Provost and president-elect of the American Council on the Teaching of Foreign Languages, addressed such responses in his opening remarks at the conference:

> It is my conviction that computers will not replace teachers. It is also my conviction that teachers who use computers will replace teachers who do not use them. This will happen just as teachers today using tape recorders have replaced teachers who do not use tape recorders. The major question is not whether computers will be used; it is how computers will be used.

Pedagogy, training, software, and hardware, as well as the implementation and management of technology in foreign language instruction, are highlighted in this issue. Two underlying themes characterize the articles: one is the clear recognition that solutions to the question of applying technology in foreign language instruction are still being formulated; the corollary theme that emerges is an appeal for practicality and flexibility.

Nina Garrett, in her article "Language Pedagogy and Effective Technology Use," discusses the application of technology to pedagogical models of foreign language instruction. She then reexamines feedback and language processing. Further, she addresses the debate about whether computers should be used for drills or creative activities and identifies training issues and teacher support needs.

Important recommendations for creating a process leading ultimately to "the seamless integration of technology throughout the whole language learning experience" are provided in Sue K. Otto's article, "Training in Instructional Technologies: Skills and Methods." Otto describes what training faculty members need in order to use technology and develop software. It is noteworthy that Otto proposes a degree program in Foreign Language Instructional Technology for Teachers and recommends specific areas of training, pointing out that curriculum coordinators and laboratory personnel should also be trained in instructional technology.

In "Effective Strategies for Implementing Language Training Technologies" Robert Ariew provides information on software tools, technologies, strategies for evaluation, and factoring-in teaching considerations. Other very useful contributions of the Ariew article include an extensive review of CALL materials, and a software evaluation questionnaire.

"Managing Innovation and Change for Instructional Technology," by Donald C. Holznagel, presents an array of practical advice based on a review of the literature and the author's experience. In this article, Holznagel introduces the term "orgware" and provides down-to-earth advice on selecting suitable managers. Next, reviewing the literature on managing technological innovation in instructional settings, he describes six areas of management concern. Finally, using his own experience as an educator, he shares advice on managing innovation and change.

The article, "Software for Language Training: Directions and Opportunities," by James P. Pusack, contains much practical information on multimedia and mass storage devices, productivity tools and authoring systems, and software/hardware compatibility. The article also provides an extensive list of recommendations for creating effective software and another for selecting, developing, and adapting software. For his analysis on the application of technology in foreign language instruction, Pusack uses four quadrants that juxtapose learning versus acquisition and productive versus receptive skills. Throughout, Pusack advocates a flexible and practical approach to software design and selection in language training.

In the concluding article, "Hardware for Language Training: Coping with Confusion," Michael D. Bush reviews principles for hardware selection, and forecasts the state of technology in terms of fidelity, feedback, and open architecture. The author's call for a paradigm shift in language learning is challenging and thought-provoking. By likening the paradigm shift to that which was caused by the advent of Copernican theory, he projects a revolutionary rather than evolutionary movement toward a new language learning paradigm.

The use of technology in foreign language education is a crucial issue for the 90s. How effectively the profession will address this issue depends not only on individual and institutional initiatives, but also on an inter-institutional willingness to share means and experience.

Applied Language Learning
1991, Vol. 2, No. 2. pp. 1-14

Language Pedagogy and Effective Technology Use

Nina Garrett
George Mason University

The topics and questions posed for this conference create an environment like that which we aim for in planning the use of technology for language students: a challenging and supportive environment for learning. The inclusion of topics dealing with the pedagogical and theoretical context for technology use is particularly significant, because these represent a crucial part of the environment that is often thought to be too familiar to need study and hence ignored. A warning footnote is to be found in a newspaper headline recaptured in the *New Yorker*: "The Use of Technology in Education is Fraud with Peril," but that should not discourage us from addressing the task.

Obviously our discussion of language pedagogy and effective technology use will hinge on our sense of what constitutes effectiveness, and how it is assessed, and that is so problematic an issue that it should be addressed first so as to establish a clear framework for the points to be made afterwards.

By now a substantial amount of research has been carried out in hopes of evaluating the effectiveness of CALL, but Dunkel's collection (1991) of papers about this research argues persuasively that much of this effectiveness research cannot, for a number of reasons, be taken at face value. For one thing, up till now there have been very few language programs in which technology was truly integrated into the overall language curriculum and into teacher lesson plans, rather than being tacked on in a more or less *ad hoc* manner. This is fully understandable. Few institutions have (or are willing to allocate to language learning) the in-house resources adequate to support the development of full-scale, in-depth course materials, or enough computers to allow their deployment for a statistically significant number of students. Outside funding is increasingly difficult to obtain, especially for materials development for the commonly taught languages—and those are the ones that draw enough students to make generalizable research possible. We find ourselves here in a kind of vicious circle: as long as an institution has too few computers to allow regular scheduling of CALL activities, teachers can hardly develop CALL materials that depend for

their efficacy on full integration into the syllabus, but as long as teachers produce materials only for optional enrichment, remedial, or motivating activities, the general perception of CALL as inherently marginal is perpetuated. This circle operates in pedagogical terms as well: as long as CALL is seen as peripheral, sometimes intriguing but not significant in its own right, it will never be truly integrated into language learning curricula, but until it is integrated, it cannot achieve pedagogical significance. And it thwarts the entire undertaking of research evaluating CALL: until CALL is integrated into language learning, research on its effectiveness is likely to be superficial, invalid, and unrevealing. But until research can convincingly support the use of CALL, it will continue to look too expensive and complicated for senior faculty and administrators to be willing to invest enough money and faculty time to achieve that integration.

Furthermore, research design finds itself in another catch-22 when comparisons are made between students' learning on the computer and off. Since the use of the computer does not constitute a method, its effectiveness cannot possibly be evaluated independently of the materials presented *on* it. If the activities engaged in by both the experimental and the control groups are the same except that one is done on the computer and one in another medium, we can hardly expect to find significant differences in the learning of the two groups. If, on the other hand, we develop materials for the computer that use its special capabilities to support quite different activities than are possible in other media—one of the most important reasons for using CALL—then it is pointless to do research comparing the effectiveness of those materials to that of ordinary lessons. Another major reason for using technology-based materials is the computer's ability to be made sensitive to learner variables and to allow different individuals quite different access to and paths through the materials—but again, research comparing the efficacy of materials which recognize and accommodate individual differences with the efficiency of conventional materials which do not is not valid.

Perhaps most important, all the studies of the effectiveness of CALL assess "effectiveness" in terms of short-term "mastery" of discrete-point, disconnected, and usually relatively trivial items of language, precisely the kind of testing that modern language pedagogy tries to avoid. It should by now be clear that we can learn very little of significance by running some students through a computer lesson and running a control group through the same materials on paper, or contrasting one isolated pedagogical feature of CALL with another, aggregating data on "mastery" across groups, in short studies that bear no relation to the principles and practices of the context pedagogical program. Effectiveness research is definitely fraught with peril.

With that as a preliminary caveat, let us turn now to the questions this presentation was asked to address. First, what theories or models have been formulated to guide the development and implementation of technology for language training? Investigations along these lines are influenced more by axioms, assumptions, and rationales than by "theories" in any strict sense, but that perhaps is inevitable at this early stage. In listing the prevalent notions or maxims that function as guidelines, two groups of "theoretical" guidelines can be distinguished. The first group of guidelines is based on *general* pedagogical mandates; the second group uses *language* pedagogy as a basis. Unfortunately, neither group of "principles" is robust enough to answer our needs.

There is disconcertingly little agreement or coherence in the general pedagogical mandates. Some writers recommend that computers should be used only for activities which cannot be carried out by teachers or written materials—others urge that computers be made to behave, as much as possible, like live, communicatively interactive, teachers. Some see the computer's greatest contribution as tutoring or drilling and handling homework so as to free the teacher's class time for spontaneous interpersonal activities, while others reject drills entirely, as inappropriate for teacher or computer, and see technology's greatest usefulness as a medium for spontaneous playful or problem-solving activities. There are plenty of educators, though they tend not to publish articles on the subject, who feel that all learning is an interpersonal activity where machines have no place at all. A variant on this theme is found in the recommendation that the computer can best serve as a stimulus for pair work or small group activities, by supplying some problem which learners discuss among themselves, i.e., with real people. Some insist that the computer should only respond to students' initiatives, answer their questions, follow their lead, not impose control on their learning. Others believe that students tend not to know how best to organize their own learning, and that computers can analyze learners' production and push them to learn more, and more efficiently, than they could on their own.

Some of our most strongly held beliefs about the advantages of computer-assisted learning must be challenged, as for example the idea of immediate feedback. It is a commonplace "truth" that even if students are doing exercises which are straight out of the workbook, at least they get immediate feedback from the computer and can therefore be assumed to learn better or more efficiently, can correct their misconceptions better, than when they get back red-inked papers two or three days later. But some research findings (Garrett, 1982) have contradictory implications: there is a very strong tendency for learners to use error feedback only to toggle a mental switch in the "Binary Correction Strategy," i.e., the students' tendency to react to an error

message with "Well, if it isn't X then it must be Y." Whether or not they are required to correct their first attempt, they clearly cannot be assumed to learn anything substantive when they simply apply this strategy. So—how valuable is immediate feedback? Or rather, on the basis of what theory can we decide what kind of feedback would actually help learning?

Another maxim of technological lesson design is based on the intuitively obvious notion that students learn best when they are interested and involved in the material. However, some teachers have found that when students are too absorbed in and enthused about the activity, they get so carried away that they cannot tolerate the constraints imposed by sticking to the target language and begin talking to each other about the task or game in the native language.

Still another favorite conviction of liberal education is that students learn better, enjoy learning more, and retain what they have learned longer, when they control their own learning, following their own interests rather than a prescribed path through hypertext or hypermedia materials. That may be true in general, or true up to a point, but the choice of materials offered in the hyper-space, and the cross-references made available among these, tend to reflect language teachers' often idealized notions of what would have attracted them as learners, or what learners ought to find interesting and relevant. Close attention to learners' actual uses of such materials is likely to reveal that often they either proceed in a narrowly linear way, neglecting to look for the tidbits of enrichment or the fascinating cross-references and connections so carefully built in for them, or they access so many of the available branches that they become hopelessly lost and distracted from the original intent of the exploration. As Fletcher pointed out (see note below), most students do not explore available options. If the teacher or the syllabus does not structure their approach to the activity, give it some purpose, students are not likely to be able to make effective use of the materials. (However, we need to think carefully about the assumptions on which we develop such structure, purpose, and guidelines for use. We should not take it for granted that the learning strategies familiar to teachers and commonly found in learner-advice manuals are based on research-based understanding of second language acquisition, and they may not be relevant at all to technology-based language learning.)

It seems that all of these recommendations have merit in some context, but that none of them should be elevated to the rank of a general guideline. None of them represent testable hypotheses; research projects cannot be devised to prove one side or another of these conflicting opinions right or wrong. In other words, none of them represents a theory.

If we look to language pedagogy for guidelines, we find strong and conflicting opinions, but not much coherent theory. Methodologists who insist on the need for explicit attention to grammar maintain that computers are ideal as drill masters, but those who reject explicit grammar teaching altogether sometimes also reject computers precisely because they believe that that is all the computer can do. Those who believe that conscious learning cannot turn into the kind of communicative ability which underlies spontaneous natural speech tend to feel that only if technology presents authentic data without any pedagogical manipulation is it worth anything to learners. In contrast, those who believe that careful structuring, explanation, and practice underlie the internalization of lexicon, idiom, and grammar (that is, those who believe that "learning" does turn into "acquisition") tend to see the computer's best role as that of endlessly patient tutor. It is a major axiom of one school of language pedagogy that adult classroom second language learning should be structured to resemble closely the natural learning environment of children learning their first language—which might seem a reason to rule out the use of technology entirely. On the other hand, this replication of the toddler's learning environment should, it is claimed, allow a sizable silent period during which students receive input but are not required to respond by producing language, and that pedagogical principle could make a strong case for highly interactive technology-based language materials which elicit responses indicating comprehension and tracking interest, but no spontaneous language production. (However, Natural Approach proponents have not argued for that possibility.)

Consider the hallowed dictum that students learn to do what they are given the opportunity to do, which was based on the recognition which motivated the audio-lingual revolution after the war, that students will not learn to speak a language if all they ever do is analyze, read, and translate it. It seems positively simple-minded to insist that of course they learn to read by reading, to speak by speaking, and so on. That principle is the basis for efforts to extend the "four skills plus culture" approach which governs much of our pedagogy to the uses of technology. Since the computer can with great efficiency present reading texts and the tools needed to decode vocabulary and grammar, it is assumed that students can learn to read using the computer the same way they learn to read using books. Unfortunately this way of thinking perpetuates our least sophisticated approaches to the teaching of reading, and ignores the computer's ability to offer learners ways of learning to read that go far beyond simple decoding. Recent research as well as ample teacher experience suggests that students do not in fact learn to read very well by decoding texts with the aid of glossaries, no matter whether the activity is carried out on a computer or the printed page, and that the activity of reading includes subtle and complex skills

like inferencing, tracking the progression of old and new information, scanning, inducing a contextual framework—all of which are extremely difficult to teach on the printed page but which allow the dynamic visual capabilities of the computer full play.

Conversely, it is assumed that since computers are not good at communicating spontaneously and authentically with learners, students can learn nothing about speaking by working with a computer—or that since interactive video delivers lively dynamic authentic visual and aural representations of culturally embedded interactions and experience, the cultural component of communicative competence is well served by technology. And so on.

However, arguments have been raised against separating language learning thus into four distinct skills. The above-mentioned "natural approach" insistence on the importance of the silent period as preparation for speaking in native-language-learning children surely suggests that listening is a crucial part of learning to speak. A wealth of parental experience, corroborated by educational psychologists, shows that listening is also a crucial part of learning to read, and that in fact all three of these skills are interlocked. Children who are regularly read aloud to, following the text and the pictures as they listen, frequently learn to read before they have any instruction at all. And indeed, a highly provocative study by Patsy Lightbown (1990) indicates that child foreign language learners not only learned to read surprisingly well when they were allowed to work on their own with no input except printed stories with accompanying audio tapes, but also seemed to learn to speak by this method nearly as well as children who had had teacher-guided practice in speaking. It might be well worth our effort to try that simple technology, or to enhance it on a computer with interactive audio and/or video, to help our learners learn both to read and to prepare for speaking. Even those who do believe that adult L2 classroom learning should be structured to replicate child L1 non-classroom learning (a position with which this writer does not concur) might acknowledge the potential to address these skills simultaneously.

As for our strong intuition that interactive video is the window into a broadening cultural experience for students, Howard Nostrand (1989) and Claire Kramsch (1989) have pointed out that unmediated authentic materials, whether text or video, are almost certain to be perceived by learners through the lens of their own unconscious native culture assumptions and preconceptions, so that those dynamic materials may in fact actively reinforce their stereotypes and prejudices.

So it cannot be taken for granted that there is any clear connection between particular language learning goals and "commonsense" ideas about how technology can help students reach them. Neither general-pedagogy nor language-pedagogy guidelines seem adequate.

The essential prerequisite to arriving at useful principles is the realization that the problem at hand is not one of language teaching, nor yet (with all respect for those who used the word in setting up this conference) of language training, but of language learning. It is pleasant to believe that learning is the natural and inevitable outcome of teaching, but, alas, all too often it is not. The real problem is that very little is known about the process of adult classroom language learning, what actually goes on in the learners' minds. Research on that process is the stuff of second language acquisition (SLA), but for a number of reasons, little of the available SLA research should be taken to have implications for the adult foreign language classroom. First, most of that research is either on children, on learning outside the classroom, or on the learning of English by groups of learners from a wide variety of native languages, or a combination of these. Second, of those studies undertaken on classroom foreign language learning, most are not really theoretically motivated, open-minded investigations of learning but instead study the outcome of pedagogical manipulations in order to confirm ideological preconceptions about what is or is not good teaching. This kind of study is really pseudo-theoretical. So little is known about how language learning actually happens that experiments on "what kind of teaching techniques make it happen better" are not yet warranted—nor is it clear what constitutes a valid measure of "better."

My emphasis on the word "theoretical" may draw immediate objections from language teachers that theoretical linguistics and the SLA research it supports do not provide insight into real learners' problems with real language data in real-time real-world learning. That is true, but the conventional assumption that the only alternative basis for practice is "eclectic" experience has prevented close examination of the development of other branches of SLA theory which bear directly on these problems. The notion of communicative competence, which is generally considered the dominant "theory" of language learning, is actually nothing more than a statement of the agreed-upon goal of language learning. It makes a well-founded claim that knowledge of language includes intuitive understanding of its sociolinguistic, pragmatic, and discourse appropriateness as well as of its grammatical correctness, but it makes no testable claims about how individual learners actually connect language meaning and language form in comprehension and production, let alone how they can do it on the basis of classroom input.

But the "nattering nabob of negativism" approach to these conventional rationales for technology use should give way to some positive statement about theoretical precepts which can be used as guidelines. One theoretical model that does lend itself well to present purposes is that of interlanguage (Selinker, 1972). Although the term is often used without theoretical connotations simply to refer to the

incorrect version of a target language spoken by learners, it does represent a testable and significant hypothesis that learner language is systematic in its own right. That is to say, even though learners' versions of the target language are fragmentary, unstable, and largely "incorrect," they do represent learners' own idiosyncratic hypotheses about how that language works. However, to describe interlanguage in terms of its erroneous deviation from the target language (e.g., as simplification) or in terms of its inclusion of native language rules (e.g.,transfer) is to focus entirely on the language product, the learners' utterances. The real focus of interlanguage research on SLA should instead be on language processing, on the individual idiosyncratic details of learners' actual real-time comprehension and production. SLA research may have as its purpose the improvement of language pedagogy or a contribution to SLA theory, or both simultaneously. This focus on the process rather than the product invokes psycholinguistic rather than linguistic theory. The study of language processing and language learning is not per se the same thing as the study of the nature of language. Obviously these must in any final analysis be very closely connected, but they are not indistinguishable, or simply two sides of the same coin. Linguists' concern with language acquisition is usually on the properties of language systems that make it possible for them to be acquired, while psycholinguists concentrate more directly on the process of acquisition in real people in real time. Interlanguage research can thus use the formalisms and methodology of work in psycholinguistics to contribute insight into the problems of language education, and that theoretical model can therefore serve as a conceptual basis for considering other issues about the implementation of technology.

How can technology best be integrated into the overall language curriculum and into teacher lesson plans? The appropriate answer, of course, is "cautiously." It is, of course, possible for teachers to assign the computer materials provided along with many foreign language textbooks for use in the language lab without making any reference to them in class activities or simply checking off that they were completed, thus not really integrating them at all. This kind of technology use may not be worthless, but the benefit realized is likely to be extremely small, since students will quickly realize that this work "doesn't count" or duplicates other efforts. Even when teachers want to make principled use of software, practical problems such as limited numbers of workstations and/or scheduling constraints often make it impossible to integrate substantive use of technology into a whole curriculum at once. The single most important component in the success of the effort.is teacher support in two senses—both support by the teacher of the students' computer-based work, and support for the teachers who want either to develop their own software or use purchased materials or in-house ones created by development teams. Even if only gradually and in

small segments, computer use must become an integrated part of the curriculum rather than an exotic activity separate from "regular" work if it is to have any significant effect.

What training and support do teachers need, in order to be able to integrate technology use? First, they need to know what information and activities the software presents, how it was intended to be used, and how to evaluate it from numerous perspectives. Does it accompany their textbook, or is it free-standing? If the latter, does the vocabulary used as context for a particular topic or at a particular point in the sequence correspond to what students know at this point in the syllabus? Is the material too generic or too abstract for students to be able to integrate it into what they are learning? Was the software designed on the basis (explicit or implicit) of a particular approach or ideology or method? This is especially problematic in regard to underlying biases about grammar teaching. Does the approach match the teacher's? Is the use of the software optional for students? Is it for enrichment and extra depth, or for remediation? Will it get (or let) students into something they have not been prepared for?

Teachers need to have a clear description of the relationship of the technology use to ongoing classroom activities. Is the software to be used in class? A single computer or videotape player with a projector can be used to present material for discussion, to demonstrate or explain new material, to provide a stimulus for student response in games or problem-solving, to follow up on homework assignments, etc. If this is to take place, who brings the machinery to class, sets it up, runs it, and returns it? Does this take place between classes, or does it interfere with class time? Does the equipment have to be reserved? How far ahead of time? Does the classroom have adequate light, window blinds, wiring, movable seats, soundproofing? Has the teacher received any training on how to manage smoothly? If several students must share a computer or interactive video workstation, class activities must be structured quite differently than if each has a machine, or if there is only one machine under the teacher's control. If the technology is to be used by students on their own outside of class, are the hours of access adequate? Are there trained monitors on duty to help those unfamiliar with technology? Is the material intended to be used in anticipation of classwork, as when students are asked to prepare a reading assignment or view some video segments in preparation for discussion, or to work through some grammar concepts or learn some assigned vocabulary? Or does the assigned technology use follow up on classwork, as in doing homework on a topic explained in class? Is it handled differently in each of these cases?

If new technology-based materials are being developed by an in-house support facility, teachers must be involved in establishing the priorities and in the development process even if they are not

themselves responsible for actual production, because only thus are they likely to feel comfortable about integrating the software into their curricula. In addition, the support unit must include professional expertise to follow up on the implementation of technology with a variety of research studies, including in-depth queries on how teachers and students feel about the technology use, how each group thinks it adds to or distracts from learning.

The most important requirement is that teachers and students alike should have a clear and purposeful idea of how each use of technology is expected to contribute to language learning, and that, of course, begins with a clear notion of how language learning happens. For this reason learners too should have some introduction (though brief and non-technical) to the concept of interlanguage; they generally find it both encouraging and actually helpful.

In designing or evaluating feedback routines, for example, teachers and designers should keep in mind the interlanguage hypothesis that learners are constantly testing their own fuzzy hypotheses about how the target language really works. Feedback can be of several kinds: a bare rejection of student input, an indication of where the error is located, an indication of why the response is incorrect, a hint as to the correct answer, the correct answer itself. In what kind of exercise, for learners at what level of learning, will which kind of feedback be most useful? Is it item-specific or generic? Is there some corroborative feedback to correct responses as well as to errors? (Sometimes it can be helpful to students who have guessed at their answer to know why it was right.) Do students have to type in the correct answer before they can proceed, and if so, is that genuinely useful? Are missed items recycled for review? If scores are kept, do they actually tell the learner anything? (Scores of responses in an exercise on the dative which includes three entirely different syntactic reasons for using that case will not help the learner at all unless the errors are tallied separately for the functional conditions and so reported.)

Interlanguage research underscores the great variability among learners as to the kinds of idiosyncratic rules they build as an unconscious framework for their comprehension and production, and that variability constitutes one of the most important benefits for using technology to assist language learning. The potential for real individualization of learning has as yet barely been recognized, but until we develop CALL equipped with programs to track the individual learner's progress through learning activities, we will not be able to design lessons which use his or her learning history, personal learning preferences, and cognitive style to suggest the most efficient and friendly learning strategies.

Similarly, individual learners will take very different paths through hypertext and hypermedia materials, and only a careful

correlation of their particular choices with their ability to work productively with the content afterwards will give teachers a good sense of the value of the hyper-materials. The design of the suggested structure or of the advice given students on how to use the materials, which should differ according to the level of learning and probably to some individual characteristics, should also be based on analysis of the ways in which individual learners shape their use of the environment.

One very clear implication of interlanguage theory and research is that we must take it for granted that adult learners (at least those who are motivated) are actively engaged in constructing their own sense of language and their own strategies for language learning, whether they are conscious of doing so or not. (In fact, they are often unable to articulate accurate insights into their own processing; if they are asked why they produced a certain answer, their report may be altogether at odds with the underlying regularity of their language behavior.) Their present level of cognitive maturity cannot simply be switched off or suspended; they cannot be encouraged or forced to learn language as the child native speaker does. It is crucial to their success and their interest in language learning that they are helped to understand the process itself and to take responsibility for their own learning. They need to know why they are asked to perform certain activities; they need to understand their own strategies (productive and counter-productive) and when and how to seek appropriate help to improve them. Most of all they need to understand what is really involved in learning how to read, how to write, how to listen, how to interrupt to ask information, how to paraphrase, how to fudge, how to take notes, etc. The computer is the ideal medium for instilling this kind of responsibility, because learners can come to feel comfortably in charge of their own progress, neither pushed nor held back by other students.

The implications of interlanguage theory, or any other approach to Second Language Acquisition theory cannot yet be taken as as a straightforward blueprint for the design of CALL materials, but I believe that they constitute important working guidelines. The most important contribution is the fundamental principle of interlanguage theory, that learners are constantly working out their own hypotheses about how the target language works as an expressive and communicative system. This implies that every CALL lesson should be designed to aid students in the hypothesis-forming and hypothesis-testing process and designed simultaneously to collect data on that process. These data can serve several purposes—to improve the software, to improve the teaching that provides the pedagogical context within which software use takes place, and to understand SLA itself. The software design and the data collection must form an ongoing cycle in the CALL development effort. Obviously we do not yet know enough about SLA generally to anticipate all the hypotheses that students are

likely to generate about a given language task or structure and we cannot hope to begin a CALL project by designing from scratch a set of lessons that will automatically infer and adjust to every possible hypothesis of every individual learner. We can and should begin by using whatever SLA research we can as a basis for our designs, follow through with the most extensive data collection that we can afford, analyze those data in the context of what we know about our learners in our particular pedagogical situation, and cycle the results of that research into the next iteration of software design.

As in every aspect of CALL development, here too it is clear that the software cannot substitute for intelligent human participation. We need to make both learners and teachers aware of basic notions of interlanguage theory, so that they can cooperate in the attempt to understand language learning from that perspective. Unless they conceive of it this way, none of the data collection suggested here will be of any use. The tracking of learner activity will yield thousands of data points but no useful information unless both teacher and student are prepared to ask interesting questions and look to the software tracking data for answers in the form of interesting patterns of performance. As we develop a greater understanding of learner language processing, we should be able to design more useful and efficient routines for adapting to specific learner idiosyncracies, but we can never hope to make the software on its own automatically able to provide totally accurate analysis of the individual learner's hypotheses. Nor would it be appropriate or pedagogically useful to aim for this, any more than it is appropriate for a teacher to regard learners as helpless passive recipients of authoritative outside analysis. The whole point is that learners must learn to take responsibility for their own learning.

Accordingly, the "individualization" of software is not a matter either of exhaustive second-guessing of all the possible answers that a variety of students might give, so as to be able to make the software provide specific responses to each, or of providing every conceivable strategy or path through a task so that students can choose the one they believe will suit them best. As has already been discussed here, and widely recognized in other discussions, students are only too likely either to ignore built-in options, or to use them inefficiently, or to get lost in exploring them. We need instead to use the implications of existing SLA research on individual cognitive and affective factors as a basis for designing a limited number of approaches to helping students to understand their own learning styles and preferences and how to structure their own strategies accordingly. Again, we then follow through by tracking how well those strategies work and how students feel about their learning. Being respectful of individual differences does not require that we accomodate every idiosyncracy, but only that we open-mindedly and objectively explore them for the purposes of helping students learn how to learn better.

Returning to the questions "What are the specific strengths and weaknesses of individual media/technologies in language training?" and

"How can this information be used to improve applications of technology to language training?", I have already suggested some answers in the section warning against too rigid a four-skills approach. It is potentially deadening to distinguish technological activities on the basis of compartmentalized goals—computers for grammar and reading, video for culture, audio for listening comprehension. In nearly every possible use, each technology is more interesting if linked to others, and the resultant learning is likely to be richer, more complex, and more lively. (All of these suggestions should of course be tested by research—but they have to be implemented even if only on a small scale, before they can be tested.) The use of audio for listening comprehension should often be enhanced by video, since comprehending speech most often happens with the support of vision, except on the radio and telephone. As CD ROM and CD-Interactive technologies become more common, the addition of audio to any materials will be possible, allowing grammar and reading activities to be thus greatly enhanced and in turn to contribute more directly to speaking competence. It should not be assumed, however, that CDI or DVI will make videotape obsolete and useless. Tape will always have its uses—for example to capture satellite broadcasts of current authentic materials and students' own classroom activities for later discussion. However, simply watching a tape outside of class is likely to diminish in value to a purely recreational exercise unless it is accompanied by detailed and structured ancillary materials such as are now being created for interactive video. (In class, of course, teachers who are thoroughly familiar with the contents can stop tapes, discuss incidents or perspectives, rewind, etc., and supply the interactivity which is otherwise possible only with videodisc.)

In all interactive materials, it is the design of the computer materials which controls and provides the context for the video and/or audio material which is responsible for the quality of the pedagogy. Expertly produced, beautiful, dynamic, and authentic video and audio can still make for really bad CALL. Again it must be stressed that the medium is not the method.

In conclusion it should be pointed out that many of the suggestions made here about the use of technology really apply just as cogently to language pedagogy in whatever medium or environment. Making best use of CALL will inevitably prod us to change some of our traditional assumptions about what supports good language learning generally. Using CALL does not force us to do anything very different, as too much of the pedestrian software now available makes perfectly clear. However, it does give us an opportunity to analyze many materials and activities whose value we do not otherwise question.

Note. I am grateful to Drs. William Fletcher and John Clark for the comments they made as discussants of the version of this paper given at the DLI conference, and have tried to incorporate their insights into this version.

References

Dunkel, P. (Ed). (1991). *Computer-assisted language learning and testing: Research issues and testing.* New York: Newbury House.

Garrett, N. (1982). Unpublished dissertation, University of Illinois, filed under Cornelia B. Garrett, in *In search of interlanguage: A computer-based study of second language acquisition* (in press). New York: Peter Lang.

Kramsch, C. (1989) Media materials in the language class. *Contemporary French Civilization*, XIII, 2.

Lightbown, P. (1990, October). *Getting quality input in the second/foreign language classroom.* Paper given at the conference on Text and Context: Cross-Disciplinary and Cross-Cultural Perspectives on Language Study, Cornell University.

Kramsch, C. (1989). Media materials in the language class. *Contemporary French Civilization. XIII*, 2.

Nostrand, H. (1989). Authentic texts—cultural authenticity: An editorial. *Modern Language Journal, 73*,1.

Selinker, L. (1974). Interlanguage. In J. C. Richards (Ed.), *Error analysis*, London: Longman.

Author

NINA GARRETT, Associate Professor, Department of Foreign Languages and Literatures, George Mason University, Fairfax, VA 22030. Specializations: second language acquisition theory, foreign language pedagogy, and the use of technology in language learning and language acquisition research.

Applied Language Learning
1991, Vol. 2, No. 2, pp. 15-29

Training in Instructional Technologies: Skills and Methods

Sue K. Otto
The University of Iowa

Claiming to know how to teach language teachers to use technology effectively probably ranks right up there with claiming to be able to herd a hundred cats through Times Square. At best it is a very risky assertion. Nevertheless, drawing from personal experience in technology development and management as well as from the insights gained by a profession that has for decades sought ways to enhance language learning through technology, this author will attempt to define a solid training program for sustained and effective use of technology in language instruction.

We have learned that the technology will not find its way into the curriculum by itself; thoughtful integration of technology is a process of assessing the real strengths and weaknesses of each technology, matching the technological capability to appropriate skills and tasks, gaining expertise in their use, and developing sophisticated materials that are of genuine benefit to the learner. There is a growing awareness that this will not begin to happen to a significant degree without teacher training (Curtin & Shinall, 1987; Kinnaman, 1990; Phillips, 1989).

The characteristics of a resident environment like the DLIFLC suggest that the application of technology can effectively diversify both in-class teaching and learning as well as out-of-class study. Training a group that shares a common equipment platform, common class characteristics, and a well-defined set of instructional goals considerably simplifies the training task. The result of a carefully developed training program could well be the seamless integration of technology throughout the whole language learning experience.

Training Teachers to use Technology Effectively in the Classroom

The basic scheme proposed here for training teachers to use technology is one of progressive integration. Through a four-phase program of instruction and curriculum development, a relatively small

number of personnel will effect changes in course design and content as well as catalyze changes in the teaching practices of their colleagues.

- *Phase 1:* Train a core group of teachers and curriculum coordinators across the languages in an intensive course on technologies presented within the context of target language skills.

The faculty selected for this training must be good teachers who have already shown an interest in technology. Teachers who volunteer out of idealism and the genuine desire to improve the curriculum are best, but other kinds of incentives may be needed to encourage this process across the schools. During the initial training phase, release time from teaching duties is sufficient to recruit interested trainees. (Other incentives will be discussed later in more detail.)

The training is conducted by a team of language teaching specialists with technical expertise in up-to-date computer, video, and audio systems. (See section on preparing teacher training staff for a discussion of how trainers should be prepared for this kind of work.) The course will not duplicate College of Education teacher preparation. The technologies must be presented from the perspective of their usefulness as tools in the whole process of language instruction, not as an object of study in itself or as an ancillary to "regular" instruction. Moreover, the relationship between technology and methodology must be clearly defined from the start: technology does not (and should not) dictate the methodology; rather the methodology will drive the use, if any, of technology. The critical objectives of this training phase will be to calibrate expectations for all the technologies to reasonable levels and to move faculty out of the traditional modes by which they were taught toward new ones integrating technology, preparing them to alter the language curriculum to include technology where appropriate.

As June Phillips (1989) points out, learning to implement technologies is a task complicated by the fact that teachers are most influenced by the way they themselves learned languages and by personal theories of teaching, which may be in conflict with practices they are told to adopt in their teaching. She goes on to say that changing these behaviors requires more than an afternoon's workshop and that we "must provide the kinds of pre-service and in-service opportunities that develop understanding of the underlying theories and assumptions in new systems and not just the tricks of using them (p. 314)." Indeed the initial training described below is at least a four-week intensive course (20 8-hour days) comprising formal presentations on theoretical and practical issues, hands-on familiarization with hardware and software, and microteaching with technology.

Before the program itself is described, it should be noted that such training does not break down into a series of single-technology

workshops, a familiar model at conferences and in education curricula: a computer workshop; a video workshop; an audio workshop (it could happen); an interactive videodisc workshop, etc. The primary point of reference is language instruction, with an emphasis on individual skills and how any given technology might be used in teaching them. This approach reinforces what our priorities should be in exploiting the technology.

Training and Sequence

Although the faculty may have some awareness of current technologies, the training will begin with a relatively brief introduction to the hardware that is available to them at their institution: computer, audio, video (tape, cable, satellite), multimedia (video, audio, and text under computer control). It is important to establish a common understanding of the technological base they will be expected to handle and to build an awareness of all the possibilities they should consider.

After this initial acquaintance with the hardware, trainers provide a general introduction to issues and options in instructional design and production of technology-based programs for language instruction.

In the present context of language methodology and technology this overview would include:

- matching teaching goals and technology solutions
- utilizing authentic vs. prepared materials
- accommodating different learner styles and strategies
- teacher-control vs. learner-control
- technological tools and their trade-offs of ease of use (training time), power and flexibility
- existing instructional models for computers, audio, and video: tutorials, practice, simulations and games, testing, productivity tools
- standards for software evaluation in light of pedagogical goals

Once this general background has been established, the skill areas are examined one by one. The following outline suggests the kinds of topics that might be covered in the training. It serves as a point of departure and is not a comprehensive syllabus, since the overall content and emphasis for each of the modules of a real training course would be heavily influenced by the special purposes of the institution.

Receptive Skills

Listening:

- schema theory and discourse analysis and their relevance to comprehension
- pre-view, viewing, and post-viewing activities for video-based listening comprehension
- collaborative listening activities with video texts
- audio: comprehension practice and testing without visual reinforcement
- strategies for using satellite transmission for listening practice; building skills to cope with authentic spoken language

Reading:

- interdependency of the receptive skills
- teaching reading skills through guided lessons with computer enhanced texts
- reading in a multimedia environment
- evaluation of reading levels through computerized adaptive tests
- special reading and research tasks in a multimedia environment

Productive Skills

Speaking:

- tape techniques (audio and video) for practice and for self-evaluation
- video as a stimulus for in-class conversation
- collaborative problem solving with computer simulations
- communicative language laboratory activities
- computerized pronunciation practice with visual feedback on performance-building simultaneous interpretation skills

Writing:

- productivity tools for writing and translation
- computer conferencing for interactive writing
- prompted writing packages for improvement of writing skills
- use of video/interactive video to prompt/focus composition

As each area is explored in depth, the trainee will practice using the technology as a student, observe model teaching with technologies in question, and then practice teaching the skills with the technology. Training will include techniques for teachers to enhance their own presence and effectiveness in the classroom as well as ways to deliver instructional experiences to students working alone or in small groups outside of class.

Because of the nature of the subject matter, not all the instruction will or should take place via face-to-face instruction; workshop activities, demonstrations, microteaching, and group discussion and practice can be included. Successfully exploring and gaining a real understanding of technology is largely a self- instructional task, but one that requires special support materials to guide the learner through the process.

Standard equipment and technical reference manuals are of little value in learning to use equipment for a special purpose or to approach an unfamiliar computer application. Likewise, a trainer can waste a great deal of energy talking about and even demonstrating a system or process that is never really understood by the students until they actively work through it. Several easy techniques can be applied to especially good effect in order to lead a trainee through a process that is best experienced as it is described:

- *Printed tutorials that lead the trainee step-by-step through a process.* This low-tech approach is particularly effective for self-instruction in mechanical tasks such as learning, to use all the aspects of controlling a videodisc player with a remote unit or the basic functions of a microcomputer operating system.

- *Videotaped demonstrations with supporting print materials.* Although training of this kind assumes many live demonstrations of techniques during face-to-face instruction, individual work with tapes and printed commentary prove very effective for intensive study.

- *Phantom student routines.* This is a familiar CAI technique in which sample student interactions are simulated for the person at the computer. Learners first observe the entire range of functions and commands of the system with accompanying commentary and then try them out.

Production of training materials like these represents important economies of time and energy for the overall training program, since their potential to instruct in future training sessions requires no further investment of production time or trainer energy.

Much technology training can be done with a mix of faculty from different languages. In fact, trainers should have the teachers try materials for a language other than the one they teach. This is the best way for them to develop a feeling for the student perspective. The notable exception is the special training required for computer-based work with non-standard writing systems and the special complications these pose. This is best accomplished through special sessions with individual languages.

Realistically, this training scenario would not be endlessly repeated; indeed the scheme is devised to get the teaching staff up to speed in a comprehensive and organized fashion. It need be repeated only enough times to create a minimum base of trained people across language schools.

Ongoing, in-service training will be quite different as a level of involvement with technology is established. Courses must be offered to fit new circumstances: many faculty will have gained expertise; incoming faculty will be taught these concepts and techniques as an integral part of the standard pre-service curriculum; new materials constantly become available; technologies change. Trainers must offer shorter seminars or workshops to update faculty on innovations emerging from internal and external research and development activities. Such staff development can be delivered by the on-site training staff or by guest presenters with special expertise. It is safe to assume that these sessions will include the same core group of faculty who were trained initially, and that they will continue their involvement in working with new technology, if they have been properly supported in their past activities with technology.

Even in a resident environment with an extensive site-tailored program of in-service instruction, the institution will need to tap external channels for education and expertise. As Ray Clifford (1987) points out, the differences between academic and government foreign language programs can be complementary, and cooperative ventures can produce beneficial results. Experts and courses, available at conference workshops or in academic programs elsewhere, may serve the training purposes of faculty in a more appropriate or cost-effective manner.

- *Phase 2:* Core faculty, technology specialists, and curriculum coordinators work together to review the options for specific languages and redesign/revise the basic curriculum to integrate appropriate technology-based materials.

The next phase requires the faculty to take the first critical steps in revamping the curriculum—applying their new training to improve the course design and content with available technology-based

materials. Viewed abstractly, this process is relatively straightforward: 1) identify areas of the curriculum that can be improved; 2) review and select new media-based programs that will be suitable for the course; and 3) weave these programs into the curriculum, identifying their exact role in the continuum and designing activities that ensure a good fit with the rest of the courseware. The extent of these modifications will depend on such factors as the amount of release time core teachers are given to work on the changes, the existence of acceptable technology-based materials, and institutional budget constraints on software and hardware acquisition.

It is unrealistic to project radical, across-the-board changes in the curriculum. While it is hard to imagine that any curriculum could not be improved with effective use of video and that teacher and student productivity could not improve through the introduction of computer-based materials, the application of technology-based materials is likely to happen in stages. Most faculty will embrace a technology that is a convincing solution for their needs. Once consciousness has been raised among all faculty members through teaching in a technology-enhanced curriculum, they will watch for new materials that might be relevant to their needs. Unfortunately, the fact is that there are relatively few materials available from either commercial or non-profit sources for languages other than Spanish, French, and German. In the case of the less commonly taught languages, integration of technology is likely to depend on faculty inclination and institutional resources available for local materials development projects.

Perhaps the most important training-related issue during this phase is that any commitment to teacher training in technology must be matched by an equally strong commitment to support the process of curriculum development. In addition to release time, faculty must have personal access to equipment for evaluating and selecting materials, as well as a significant budget for the acquisition of programs that will be integrated into the curriculum.

At the same time that curricular changes are being made, plans for delivery support must be made as well. Technology-based instruction demands special equipping of classrooms for teacher use of the technology and reconfiguring or re-equipping lab space for small group and individual student work with the new technologies. In addition to hardware for the lab setting, sufficient software must be acquired to handle the demand of out-of-class assignments.

The final step in this phase is the development of orientation procedures for the rest of the faculty and staff who will be affected by the changes that have been planned.

- *Phase 3:* Implementation and expansion of expertise.

The teachers who have not been involved in the revision process must begin to take an active role in planned changes. Their training to

assume this role will be handled by their colleagues from the core group largely during class and laboratory hours. Initially, however, they will be oriented through a short "literacy" workshop on audio, video, and computer technologies within the skills context (a much abbreviated version of the training received by the core group). This workshop should be designed to focus on showing convincing technical applications in order to demystify the technology and overcome teachers' feelings of intimidation and skepticism.

Next comes familiarization with the redesigned curriculum. The core teachers discuss and model the new strategies and materials devised for the courses. The "novice" faculty, in turn, practice with the technology and techniques that are part of the new curriculum.

The transition into the new technology-based curriculum need not be an abrupt one. The changeover can be smoothed by using a team-teaching approach at the beginning that assigns a core-group member in the classroom to serve as model and advisor for the regular instructor. Later, as the regular instructors gain confidence and experience, they will play a vital role in the ongoing evaluation of the effectiveness of the materials and in any revisions that take place in subsequent courses.

- *Phase 4:* Periodic Reassessment.

Thus we complete the first cycle of re-educating faculty to break away from traditional methods of instruction. As new programs and technologies present themselves and promise to enhance our repertoire of teaching tools, we must repeat the cycle, reassessing the value of the old against the new and redesigning curriculum to better accomplish our instructional goals. In a setting where the curriculum is closely prescribed for its teaching staff, this reevaluation should ideally take place after every course and at least every two years.

Technology has always been years ahead of its thoughtful integration into the curriculum. Perhaps the biggest challenge in technology training is deciding how and when to update faculty on new technologies. We must strike a happy medium between trying to keep up with galloping technology and the constraints (time and money) of redesigning the curriculum to integrate new resources as they appear. Clearly, the philosophy of "pushing no technology before its time" applies here and the problem can be solved by relying on the guidance of the Research and Development staff. The R&D staff form the vanguard that forges ahead into new territory, bearing the frustrations of untried hardware and software, producing pilot programs, evaluating their effectiveness, and refining the best of them. On the other hand, the training staff concentrate on developing teacher use of mature technologies and robust programs that are viewed as viable by the R&D staff.

Often teachers find that materials developed elsewhere do not satisfactorily meet their needs. If their training is successful, instructors will come back to the R&D staff with ideas for modifications to existing programs as well as design specifications for new applications. This relationship will be examined in the next section.

Training Required by Teachers to Develop Effective Language Training Software

The realities of course loads, interest, and motivation lead to the basic conclusion that language teachers are not likely to become sophisticated programmers. If not as programmer, then what role can a teacher play in the development of effective language training software? The obvious response to this question is that the teacher will work as part of a development team including a programmer, an instructional designer, and a technician serving as the subject matter expert and materials designer. This is, however, not a role a language teacher can play without special training and experience, primarily with computers. Although print, audio, video, and computer are still considered (and delivered) as separate resources, all these technologies have in fact begun to merge with computer technology. An intimate understanding of computers will therefore be essential to designing the new generation of language learning materials.

The training, curriculum revision, and classroom teaching experience discussed in the first section of this article constitute the prerequisites for the course one must follow to become an effective software developer. Not all teachers who use technology can become good software developers, however. Even given substantial training and experience, a person must also have special aptitudes for design and detail, imagination, and energy to be a successful developer.

In this phase of training, teacher-developers will expand their general knowledge of instructional computing and venture into the practical production of software. Like the initial training course, such training is not accomplished with a couple of short workshop sessions. The process of developing this kind of expertise can easily spread over years; but much of the basic groundwork could be laid in 60 hours of face-to-face instruction with outside time for hands-on development practice. The expected outcome of this training is to provide the teacher enough background to be able to imagine realistic, complex applications, to write reasonable specifications for a programmer, and to program prototype interactions.

Although the object of the training is not to produce a sophisticated programmer, the teacher-developer must acquire some programming experience to better understand the possibilities and limitations of the systems being programmed. This course provides training in three different modes of authoring/programming:

1) authoring of exercise materials with template programs designed for language learning;

2) designing and programming a simple task with a

standard programming language;

3) designing and programming a moderately complex lesson with a sophisticated authoring system that supports multimedia interactions.

As programming skills are being developed, a number of the key concepts and issues introduced in the first stage of training are treated in more depth:

- authoring tools specifically for language instruction vs. generic authoring tools
- potential of digitized audio and video to enhance CAI
- problems in answer-judging and natural language processing
- design considerations for the multimedia environment
- the implications of current language theory and research on program design

Just as training teaching staff to use technology implies other institutional commitments of release time for course revision and equipment for evaluation and delivery, development training implies specific kinds of support for developers. Significant materials development projects require long-term investments of time by teacher-developers and of personnel and equipment resources by the institution.

Development time for technology-based projects will depend on their scope and complexity. Major projects progress through a number of orderly stages:

- general design of project; acquisition of development hardware and software
- elaboration of specifications for programming and subject matter base
- preliminary testing (formative evaluation) of materials
- refinements and further testing
- packaging (including development of documentation)

Excellent programs do not materialize from teachers working evenings and weekends at the expense of teaching or personal commitments. To optimize project efficiency, teacher-developers should ideally have full release time from teaching duties through the end of the period of creating the programming and the subject matter base as well as partial release time for the balance of the project.

Even given release time, faculty must have additional incentives to engage in long-term development activities. In the academic setting we are familiar with the rewards that drive all professional activities: promotion, tenure, and salary increases. In any teaching setting, rewards in the form of additional salary or increased job security will

encourage many teachers to become involved in serious development work.

The model of lone developer—or even the single language teacher working with the technical team—is only one development approach, and not always the best one. The development process can be enriched through the collaboration of two or three language colleagues working together. Several minds critiquing work in progress or brainstorming to invent new interactions invariably produce more complete, more balanced products than is possible with a single perspective.

Development of new programs is not cheap. Nevertheless, the institution that encourages its faculty members to become developers can also promote certain economies that will enhance the benefits derived from project work. The more language professionals can develop generic tools and tasks for technology-based systems, the greater the extent to which technology will make inroads into the curriculum. Developers should be trained to conceptualize at an abstract level that emphasizes language objectives over specific language content. They should be encouraged to maximize the flexibility of any system they develop to accommodate data other than that which is being generated by the specific texts chosen for the initial development. The resulting templates and models can then be used for more than one lesson by more than one language to produce materials. Collaboration and communication among faculty across departmental or school lines is essential to this kind of development.

Magnitude is also an important consideration in development activities. Although there is much to be said for massive programs such as elaborate simulations and intelligent tutorial systems that deliver instruction on many levels to all kinds of students, there is room for humbler projects as well. Not only is it unlikely that many faculty can be released for extended periods, but few will have the aptitude and perseverance it takes to direct long-term complex projects. We are still exploring strategies and techniques for computer-based learning; and many of the best contributions to the field of technology-mediated instruction come from a single good solution to a specific need. Some combination of large and small projects is needed to produce an overall impact on the curriculum; and good development programs need to be able to accommodate both broad and narrow initiatives by its faculty-developers.

Training Required by Other Staff in a Language Training Institution

The technology training requirements for a language teaching organization—whether it be a federal government school or in an academic program—are complex. Different kinds of training are needed for a number of purposes: staff trainers are needed to train the teachers to be users and developers, others are needed to train non-professional language lab staff, and students will be trained both by their teachers and by lab staff both in and out of class. How should these trainers be prepared for their jobs?

Teacher Training Staff

The training courses proposed in earlier sections of this document are handled by a professional teacher training team, whose members deliver much of the training on site, recruit guest presenters for workshops, and arrange for external educational opportunities, including coursework at other institutions and workshops sponsored by professional organizations. Looking back at the content of the technology training courses gives an excellent indication of the training needed to prepare them.

Obviously, these people should be graduates of a leading PhD program in Foreign Language Instructional Technology Training for Teachers. The prerequisites for this degree are an undergraduate degree in foreign language with an emphasis in foreign language methodology and other coursework required for a teaching certificate, and several year's experience of classroom teaching of foreign languages. The program would prepare them in the following areas:

- Curriculum development for foreign language courseware
- Instructional design and strategies for foreign language courseware
- Survey of computer applications for foreign language instruction
- Language laboratory methods
- Using video in the foreign language classroom
- CAI languages
- Computer programming and natural language processing
- Computer-based simulations and tutorials for language study
- Theories of second language acquisition
- Production of audio and video
- Evaluation and testing of foreign language skills
- Basic statistics and research methods

Unfortunately, no schools offer such a degree program; most of the people who work in this new field have been trained through on-the-job retooling. This situation is not likely to change until academic preparation courses for language teachers recognize applications of technology as an essential part of methodology training. Under the best of circumstances, the technology training that language teachers receive comprises a few superficial modules in CAI or video as part of a methodology course or a few generic College of Education courses in technology.

For the time being, the best solution is to hire language professionals who have some additional education or experience in instructional technology. The institution must then afford them opportunities to complete their training in computing, CAI development, instructional design, evaluation and research through additional coursework and self-instruction. The importance of having practical knowledge about language teaching that is a discipline-specific context rather than a standard College of Education preparation for the technology cannot be overstated. The worst ideas for development projects come from technology and instructional design experts with the meager background gained by their experiences in beginning language classes in high school or college. Such experts suggest projects such as producing a videodisc of footage taken in different stores in downtown Iowa City, with one soundtrack in Spanish and the other in Japanese. It is easier to train a language expert in technology than to train an instructional designer or CAI expert in language pedagogy.

Curriculum Coordinators

The training courses described here have concentrated on educating the classroom teaching staff. However, in institutions that have curriculum coordinators whose primary duty is to design and manage the curriculum, the coordinators must have as much training as the in-class teaching staff. They must have equal expertise in the technology so that they can collaborate with faculty from the core group to develop sound technology-based curricula.

Laboratory Personnel

Language students are rarely on their own with the technology. In most language learning programs, students spend time working with media-based materials in a language laboratory setting. Not surprisingly, the kind of training required for these people is quite different from that required of teachers who wish to use technology. In most settings there are both professional and non-professional staff in the media center, and their primary role is to facilitate delivery of

technology-based instruction. The professional staff administers the center's library of materials, equipment, and non-professional personnel. The technical training required by the modern multimedia lab professional may include computers or local area networks, media editing and production, and design and maintenance of media (audio, video, computer, multimedia) facilities for individual, small group, and whole-class use.

The non-professionals generally "monitor" the facilities, checking out materials, troubleshooting problems with equipment, and helping students who encounter difficulty with the programs. In some cases, these lab assistants are the ones who actually train students how to use the equipment and programs (though this may be a task best performed by the teachers themselves). Lab assistants must be trained to perform routine maintenance and troubleshooting of equipment failures as well as to provide basic consulting to students who are having difficulty getting started or working through the tasks assigned them.

Summary

Technology training for language teachers has few, if any, established models. Until training in the use and development of technology-based language learning materials becomes an integrated part of academic teacher-training programs, we will have to rely on special pre-service and in-service training courses to provide the background necessary for exploiting the full potential that today's technologies offer to language learning. Of course, the most problematic issue in establishing such training courses is finding enough qualified trainers. Some language professionals have found ways to gain special expertise in technology-based learning; their help must be enlisted to build training programs for teachers.

In a large language instruction program, technical training should, for practical reasons, start with a core group of good teachers who are enthusiastic about technology. Provided with intensive instruction and practical experience in using media-based materials, this group will revise curricula to exploit the technology and then train their colleagues to use the technology through modeling and coaching. Once a firm base of expertise has been established, this process is perpetuated through a pre-service program with a fully-integrated technology dimension complemented by a program of shorter seminars and workshops designed to keep all faculty up to date.

The natural extension of training teachers to use technology is training them to develop new materials. In this phase of training, teachers develop expertise in computer authoring, instructional design of CAI materials for language learning, and multimedia systems. They are then prepared to join development teams of programmers, instructional designers, technicians, and other colleagues to create new materials tailored to their specific curricular needs.

The institution must be prepared to support all these training initiatives with additional major commitments of faculty release time; hardware and software for development; research and development personnel; and appropriate facilities (hardware, software and staff) for delivery of the enhanced curriculum. Without these resources, the training will have little impact on teaching.

References

Clifford, R. T. (1987). Language teaching in the federal government: a personal perspective. In R. D. Lambert (Ed.), Foreign Language Instruction: A National Agenda, (Vol. 490, pp. 137-146), *The Annals of the American Academy of Political and Social Science*. Newbury Park: Sage Publications.

Curtin, C. O. & Shinall, S. L. (1987). Teacher training for CALL and its implications. In W. F. Smith (Ed.), *Modern media in foreign language education: Theory and implementation*. Lincolnwood, IL: National Textbook.

Kinnaman, D. E. (1990). Staff development: How to build your winning team. *Technology and Learning, 11*, 2.

Phillips, J. K. (1989). Challenging teachers and harnessing technology (pp. 312-319). In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics, (Language teaching, testing, and technology: Lessons from the past with a view toward the future.)* Washington, D.C.: Georgetown University Press.

Author

SUE K. OTTO, Director of the Language Media Center and Adjunct Professor of Spanish, University of Iowa. 125 Schaeffer Hall, Iowa City, IA 52242. Specializations: instructional technology, development and programming of foreign language computer-assisted instruction (CAI), videodiscs and interactive videodisc software.

Applied Language Learning
1991, Vol.2, No. 2, pp. 31-44

Effective Strategies for Implementing Language Training Technologies

Robert Ariew
University of Arizona

Technology continues to make an impact in teaching foreign languages. However, implementing technology in the classroom often is a confusing and complicated enterprise. In this article I will discuss strategies of implementation, looking first at authoring systems and selecting software, then at pedagogical strategies in computer-assisted language learning (CALL) and finally at the problem of evaluating software.

Software Tools

During the prehistory of CALL most people's first encounter with a computer was in trying to program the beast by means of a computer language such as BASIC. Only the stouthearted ever tried any lower-level languages, such as machine language or assembly language. BASIC was (and still is) a multipurpose environment, one that affords the user a wide flexibility. But the users who wanted to write computer-assisted language learning (CALL) materials found themselves in need of special-purpose routines, such as sophisticated text display tools (to display foreign accents or to illustrate with graphics) and more complex string processing (to accept student answers and mark them up). That is why many writers of CALL materials soon migrated from BASIC to applications languages such as Pilot. They found that these programs gave them the necessary tutorial procedures with which to build CALL materials. At the same time, users of application languages realized a loss in flexibility and speed of execution when compared to high-level languages. That was the trade-off that was implicit in the choice: increased ease of use, more tools, for less flexibility.

Often the trade-offs were acceptable. Some users found that the application languages were exactly what they needed, since, although more restricted, the procedures available in the application languages fit their instructional strategies. They argued that "if these tutorial

procedures are consistent with the educational strategies that we wish to employ in our educational materials then we have gained considerable ease of effort without serious loss." (Merrill, 1985, p. 90). However, some CALL authors wanted to go further.

Authors who took one more step toward efficiency by using an authoring system usually found that they are even more restricted. Typically, authoring systems do not require programming experience but do encourage and promote a specific instructional strategy. Again, if this strategy is acceptable for the author's needs, then using the authoring system provides considerable gains in development time.

The tradeoffs, as the CALL authors went up the scale from machine language to assembly language to application languages to authoring systems, involved less flexibility and lower execution speed, but a potential savings in development time. In the continuum of tools, as they become more specific to CAI or to CALL, they become easier to use and less flexible. On the other hand, the more flexible general purpose languages are difficult to learn and implement.

More Modern Tools

More recently, while first generation authoring systems are still in use, a new category of software tools is emerging which is having a considerable impact on CALL software development. These tools involve the concept of "Object-Oriented Programming" (OOP). In object programming, a software routine can be written to include characteristics which make it behave more like an object rather than a piece of program. The code may embody characteristics such as object size, behavior patterns, color, and other features. Once defined, the object exhibits its characteristics in a variety of environments. For example, one can imagine a "button" on the screen which causes a certain sequence of events to begin once the user "pushes" it with a mouse, by clicking on it. The button could cause an animation sequence to begin or digitized audio to be played. The button itself acts like an object: one can move it around the screen, label it, cut and paste it in various places, duplicate it, etc. Wherever it is moved, it still retains its original characteristics, namely to begin the animation sequence or to play the digitized audio. Whenever it is duplicated, each of its clones retains the functions that were defined for the button (Goodman, 1987).

OOP has implications for more modern authoring systems. Once objects are properly defined, authors can manipulate their features without having to take care of all of the programming details involved. Authors can also see what will appear on the screen and how the user (the student) will experience the software. The process of designing the

materials becomes a part of using them. One may even begin to consider that CALL materials are being produced without programming involved (although fundamentally, the process would still be called programming). The task of the author is no longer to deal with the details that make the beast behave, but to provide it with information as in a data base where the information might include digitized audio, video clips, graphics, language samples, etc.

However, although the new tools imply a significant change in the author/teacher's role, some relationships do not change: 1) It is still not a trivial matter to design and implement CALL; 2) All authoring systems (OOP or otherwise) imply restrictions in teaching strategy. Still, if objects do what authors perceive as useful, or if OOP strategies are acceptable for the author's needs, then using an authoring system provides significant gains in development time. If, however, the strategy is not useful, then no gain can be realized and other means must be devised.

Selecting CALL Materials

While the issue of creating CALL software by means of an authoring system is very important, another strategy may be chosen: selecting CALL materials. Educational software publishers are relatively few in number, but they are beginning to make a variety of materials available for several different machines. It would be a great waste of time and energy to devote months or years to creating CALL materials only to find out later that similar materials have become available and can be purchased for a fraction of the cost of developing them. The foreign language teacher must therefore adopt a strategy for either *creating* or *selecting* CALL materials.

The difficulty lies in being able to predict the type of materials which publishers are likely to make available. Those are the materials which teachers should not be involved in producing. Conversely, it would be nice to know the type of materials which publishers will not be designing. Those are the materials which teachers will not find anywhere and might consider designing themselves.

Because the costs of development are high and profits from software are generally low, publishers look for markets where there are large numbers of students and consequently large potential sales. Of course, publishers must consider several other factors, including availability of hardware on which to use the software and cost of software development (including cost of graphics, sound, color, programming, etc,). But one of the most important factors is the size of the potential market (the numbers of potential buyers of the software). That is why the greatest number of software packages is available for the first and second years of the most commonly taught languages.

Software for first and second year Spanish and French is most commonly available. German and Russian are less available, while the less commonly taught languages such as Arabic, Japanese, Hebrew, Chinese, etc. are almost nonexistent. It should also be pointed out that these less commonly taught languages present some difficult CALL design problems such as displaying ideographs or right-to-left writing.

Similarly, software will probably not be developed by publishers in third and fourth years of language, even for the more commonly taught languages, simply because there are relatively small student populations there. CALL materials for first and second years of the more commonly taught languages, however, should be plentiful. Foreign language teachers should be involved with the *selection*, not the creation of materials in this area. Materials for the third and fourth years of a language may be an appropriate level for teachers to *create* their own materials. Similarly, materials for language for special purposes (Business French, Spanish for Social Sciences, etc.) will not be readily available. This is another appropriate place for foreign language teachers to design their own materials.

Media Type and Development

We have already seen that software may either be selected or designed according to whether it is (or may become) available from publishers. Software may also be too expensive to develop in-house. There are certain software projects that may be too complex, or require too much hardware, or too many skilled personnel to be feasible with the resources available in any given institution. For instance, some software projects may need a team that includes software designers, coders, testers, etc., and the participation of more than one insititution. Such large projects unfortunately also require commensurately larger amounts of funding. For example, projects involving video technology tend to be expensive and necessitate multi-institutional participation. They are difficult to undertake as small projects. For this reason, video projects tend to include materials that are selected rather than designed in-house. Often software designers who want to include video materials must wrestle with the concepts of adaptation of pre-existing materials (repurposing) and must rely on organizations that make available video images.

Software projects, therefore, must take into account the language addressed, the level of the language, the media type, and, as we will see below, the primary language skill.

Features and Limitations

It is easy to fall into a trap and think of the new technologies as embodying the solution to all the problems of teaching and learning second languages. Technologies do have limitations. For instance, there is no technology that is at this time capable of interacting with the user in a conversational oral way. Unless giant leaps are made in the field of voice recognition and output (and similar leaps in linguistics, including semantics and parsing), the computer will not be able to converse with the user in a freeform oral interchange in the near future. But freeform conversations between two people is a major goal of foreign language learning. One can argue therefore that the new technologies are ineffective in teaching foreign languages, just as one can argue that the book is an ineffective medium of instruction (since it cannot provide freeform oral interchange in a foreign language either). Technology simply does not address one of the major goals of foreign language learning: speaking.

Where, specifically, do the strengths of CALL lie? What follows is a media analysis comparing the features of texts, video, and CALL on the four skills (listening, speaking, reading, writing). In a typical learning situation, the skills are not taught in isolation. They are most often linked together, where several are addressed by the language learning activity. For instance, in a dictation, the student listens and writes. In a reading activity, there is often reading aloud and comprehension of written material, followed by a writing activity. If we consider the specific language skills one at a time, we are simplifying (and perhaps even falsifying) what happens in a typical classroom. However, by examining each separately, we are able to isolate the attributes of the technologies by their strengths. Of the four skills, which one(s) can the computer address? Where does video shine in its capability for teaching? In what areas does it have attributes that would make it valuable in foreign language learning? In a recent article (Ariew, 1987) I discuss the relative capabilities of video, CALL and text. The chart below has been modified to take into account the latest advances in technology:

	Textbook	CALL	Video
Listening	-	+?	+
Speaking	-	-	-
Reading	+	+	+?
Writing	+?	+	-

In the textbook column, one must admit that the text is primarily a medium for addressing written skills. It shines as a medium of

presentation for reading materials; it does not have any direct ability to affect oral skills. While it can be argued that as the textbook does address oral skills indirectly by providing activities and materials that will spark discussion, it is not possible for it to provide models for oral interactions (except in a written mode), nor can it directly influence pronunciation, intonation, oral discourse, discourse strategies, etc. The text gets minuses in listening and speaking. In writing, the text is used as a means to elicit written responses and to ask for participation in written activities; it has traditionally been used to teach some aspects of writing. However, the evaluation of the writing is left up to a teacher. Since it cannot provide either feedback or a conducive environment for writing (by allowing the movement of paragraphs, outlining, pointing out spelling errors, etc. as a computer can), it earns a plus with a question mark.

In the CALL column, CALL earns a plus with a question mark for listening. With recent advances in digitizing speech on a computer, CALL demonstrates an ability to present listening materials (an ability that the book does not possess). The technology, however, is still new and there are still some problems to solve. For instance, the process requires very large amounts of storage which are typically not available in microcomputers for student use (in a laboratory situation). CD ROM may solve the storage problems shortly, however. (See Lambert and Ropiequet, 1986, about CD ROM's potential; for a CALL perspective, see Woodbury, 1988). CALL earns a minus for speaking since no program can interact in a freeform oral fashion with students. (Some point out that useful work has been done on pronunciation by means of oscilloscopes and intonation analyzers. While pronunciation may be part of speaking, it is certainly a small part and does not involve communicative abilities per se.) CALL is a good medium for reading skills instruction; it should be pointed out, however, that recent research shows that people still prefer to read from a book rather than to read from a screen. Only very short messages (one or two screens long) are acceptable on a computer monitor. CALL can address writing as well. Many programs exist to help develop ideas, organize essays, check spelling, check grammar, etc. Unfortunately, CALL cannot evaluate extended writing as would an instructor. However, the holistic suggestions which software can provide may be valuable.

Video has altogether different attributes. Like CALL, it can facilitate listening skills. It is especially suited for listening since it can not only present audio material, but visual material as well, thus increasing motivation, adding realism, immediacy and important non-verbal information to the communication. While video can also address reading skills, resolution of video images is still very poor and only a few characters may be displayed on a video screen. (A computer display, on the other hand, has more resolution.) One loses legibility very quickly

when displaying large numbers of characters on a video screen. In any case, reading text from a video screen is tiring and not altogether practical. Video earns a qualified plus in reading. In speaking, video earns a minus because, while it is true that some learning may happen as a result of listening, there is no interaction with video and hence no directly attributable benefit in using video to teach speaking skills.

Several things emerge from a close look at this table. None of the three media analyzed addresses speaking specifically. The only "resource" that can directly affect the teaching of speaking skills at the present time is interaction with a teacher (teacher-student interaction). One can also see that CALL, even though it does not address speaking skills, is still the most potentially useful medium for foreign language study of the three media. It can address reading and writing, and, given the proper storage device, listening as well. In addition, it is an interactive medium and can react to student input.

We have now discussed strategies for creating and selecting CALL materials and have analyzed the attributes of several media; we turn next to a discussion of pedagogical strategies for CALL.

Just as we can appreciate the evolution of tools from BASIC to OOP authoring systems to create CALL materials, an evolution of pedagogical strategies has taken place. Drill and practice formats have given way to materials that let students browse elaborate databases to find their own information and take charge of their own learning.

New Approaches

Foreign language pedagogy has evolved in a significant way in the past ten years. Foreign language teachers are gaining new insights about how students acquire foreign languages. The proposed theories have also had a pragmatic impact in daily classroom management and techniques, triggering changes in them. CALL has also been affected in the process. New approaches for CALL design have been proposed recently that depart significantly from previous ones and which are based on more modern concepts of foreign language acquisition.

For instance, new insights come to us from Discourse Theory which posits that to learn a foreign language it is not sufficient to study grammar at the sentence level, but that one also needs to participate in communication and to use language to solve specific, well defined problems which have some meaning to the learner. Many foreign language texts now include appropriate communication opportunities in the foreign language. Students are asked not only to perform tasks at the sentence level, but also (when they are able), to employ communicative strategies.

One of the ideas that Krashen (1985) proposes in his Monitor Model is that of comprehensible input. By lowering the student's

"affective filter" (by lowering anxiety factors and increasing self-confidence), the amount of input that the student is able to process is increased. Providing language materials at a level of comprehension appropriate to the students' abilities also is a way to enhance language acquisition.

In the area of classroom techniques, the accent has recently been placed on increasing interactions; functional syllabi are showing up in text materials; the *ACTFL Proficiency Guidelines* (1986) are beginning to make an impact on language training. While older pedagogies stressed the importance of the teacher, newer ones stress the role of students as the center of the instructional process. They are taking on more of the responsibility for their education, while, at the same time, communicating among themselves more often through group activities. Teachers, on the other hand, are assuming the role of orchestrators and mediators of learning.

While some movement toward more communicative modes of language learning has been felt and is beginning to influence the classroom, the process is at its infancy. Much more needs to be done. Some argue (Schulz, 1990) that the process is not going fast enough and that texts have a long way to go before they can be called communicative.

Changes in CALL

As changes are introduced in foreign language pedagogy, similar changes are occurring in software design. The old categories of software types no longer hold entirely. New categories of student-software interactions have had to be defined. "Practice" software, although still quite prevalent, has taken on a new role.

A decade ago, CALL categories were well understood and were used in shaping lesson materials. The four "classic" categories of software were *tutorials, drill and practice, simulations* and *games*. These categories served well for quite a while and most CALL software was designed along those lines. However in 1984 Hope, sensing the inadequacy of the old typology, writes:

> In foreign language CAI, this typology quickly breaks down. Good tutorials involve extensive practice and "simulate" the classroom. Drill is often dismissed as mechanical or "Skinnerian," largely because weak programs offer the student little help in reaching a correct answer. Advanced types of structural practice should be classified as problem solving. Simulations may be little different from drill if they can handle only a small number of precise foreign language responses to a given situation; if they use multiple-choice questions, they may be better classified as reading exercises. While we use the traditional classification scheme, we recommend that it be viewed skeptically, especially when these terms are invoked in a

> judgmental fashion, in order to heap scorn upon drill, or to extol the glories of simulation. Good CAI materials in foreign languages seldom meet the challenges of the field in predictable ways (p.18).

Similarly, Ariew and Frommer (1987) write:

> Although these classifications may have sufficed when computers were first used in language learning, they no longer cover the wider range of formats made possible by improved technology and increased experience (p.179).

They go on to describe several new categories of software. Advances in CALL have also been made by using traditional software in new and creative ways. For example, drill and practice, while they are still used extensively in foreign language study, now occupy a new place in the learning process. Jones and Fortescue (1987) argue that the computer as drill-master still has a place in foreign language learning since it can free up the instructor of repetitive tasks. The instructor can then spend more time interacting with students and give them more face-to-face communicative practice.

Along the same lines, Clark (1988) in his article, "Toward a Research and Development Strategy for Computer-Assisted Language Learning," proposes that CALL should be "a part of the total instructional system which also includes the live teacher, textbooks and other print materials, outside-of-class learning opportunities, and numerous other non-technological components." (pp. 5-23)

New CALL Strategies

Mirroring recent changes in Foreign Language pedagogy, CALL is being redefined along the same lines. Rather than to propose more software that is repetitive and rigidly structured, new CALL software design entails giving students more flexibility to seek and find information, to control not only the way they learn, but also to some extent the material they learn.

For example, some new CALL materials are structured in such a way that students are given free reign to explore (or "browse") databases. Students are learning by discovery and by doing. This kind of freeform learning is very much reminiscent of some of the more modern classroom techniques such as group learning or student-centered learning. In "Hypermedia for Education in the Life Sciences," Jaffe et al. (1989) propose that the computer should now be conceived as a Personal Learning Workstation where students take more responsibility for their own learning and where the computer provides as much or as little information as requested. Just as teachers' roles have evolved from leaders to facilitators of instruction, computers can

now be seen as facilitators of instruction. Instead of limiting the use of a computer to a one-to-one setting, teachers are also using them as information providers for small groups. This too can be seen as an adaptation of CALL design to recent advances of foreign language pedagogy. Young (1988) is more categorical when he proposes that CALL will make its greatest impact in the area of learner interaction.

Software based on student-directed interactions is coming to the fore. For example, a whole new category of CALL materials is being developed: *tools*. This category describes the use of the computer as a resource, as, for example, when students use the microcomputer to solve a broad problem. In this case, the machine itself does not hold the solution to the problem, but is a means to solve it. It is to be contrasted with other categories where the computer not only holds the answers but also the learning strategies which students adopt when they use the software. The clearest and most often cited example of tool software is using the microcomputer for writing. Typically, pre-writing software (tutorial software) helps the student define the topic, while word processing software (tool software) makes the editing and reediting cycle much easier. In addition, help is available for the student to check certain aspects of grammar (see grammar checkers), spelling (many word processors now include spelling checkers which are available in several languages), vocabulary (thesauruses are now also available), conjugation and grammatical rules (databases with grammatical rules are now becoming available, even for foreign languages), as well as idiomatic language organized according to language functions.

Another type of student-directed CALL software is *hypertext*. Elaborated in the 60s as a system to link large amounts of information, hypertext has the potential to make a strong impact on CALL. While few hypertext applications for CALL exist today, the examples point to interesting materials. For instance, the words in a foreign language reading text have been linked to information that provides explanation about the language (vocabulary, tenses, idioms, etc.) The information could just as well be linked to details about the author's life, to a dictionary, or to comments about the cultural context (politics, life styles, etc.), or even to comments about the meaning of the text, to the gist of paragraphs, etc. If the hypertext concept is taken to the next logical step and incorporates other media, there could be links to an audio recording of the text, to graphics giving information about cultural conditions, etc., and even links to a video recording, showing the specific setting or even actors playing the roles described in the text. This extension of the hypertext concept, called hypermedia, is an important conceptual leap and will impact on the design of CALL software.

Strategies for Evaluation

Once in hand, how should software be evaluated? Hubbard suggests (1987, p. 227) that the effectiveness of software is difficult to assess. He makes his argument in several ways. First, he posits that several factors inherent in software make it more difficult than texts to judge. Among those factors:

a. Software evaluations have made their appearance only relatively recently, while reviewers of textual materials have had many years experience,

b. One can't glance at software as one does with texts to get an overview of the materials. Typically software must be looked at in a lockstep fashion, selecting parts of it one at a time. Quick overviews are difficult, if not impossible to do.

c. Lesson components are quite complex, requiring several passes through the material to view all of the possible branching sequences and feedbacks.

d. Software adds several dimensions of evaluation when compared with texts. The visual and auditory aspects of software must also be evaluated as part of the review process.

e. The process of evaluation must also account for the interactions in the software. Appropriateness and quality of feedback must also be assessed.

f. Finally, there exists a hardware dimension to software evaluation. One must take into account the availability and features of the machine required to use the software. All these factors, when taken together, make software evaluation a difficult proposition.

Factoring in Teaching Considerations

Hubbard further states that "general-use forms and questionnaires [for software evaluation] ... do not focus on the relative fit of the software to instructional approach and require supplementation to extract meaningful judgements whether the materials are likely to accomplish their advertised ends." (p. 229). He argues convincingly that software evaluation must take into account language learning approaches, learner strategy orientations and type of syllabus (structural, , situational, notional-functional, or content-

centered). At the end of his article he proposes a supplementary evaluation form that takes into account these teaching considerations, as well as a step-by-step plan for software evaluation.

There is no doubt that the evaluative process must take into account Hubbard's teaching considerations for they are an integral part of the material and its quality. It is impossible to decide on software quality without first understanding the language learning and teaching goals embodied in it. However, there remain a few additional difficulties: With much software it is impossible to tell exactly what language learning approach it takes since it may take more than one. For example, it is possible to see behaviorist as well as explicit learning approaches in a single software package. Similarly, the same materials may begin with a structural syllabus and progress to a situational or functional syllabus. Software that is eclectic in nature is rather common. The evaluator may become perplexed and believe that the task of evaluation is impossible.

There exist several software evaluation questionnaires (see for instance the one used by CALICO in the *CALICO Journal*). Without adding to those existing, we propose four very broad categories of questions that may be relevant to software evaluation:

Goals

- Are the goals worthwhile?
- To what extent are they reached?

Clarity

- Are explanations on use of the software clear?
- Are explanations on content clear?
- Is there an adequate number of examples?

Quality

- Assess the quality of the language used (accuracy, authenticity, etc.).
- Assess the programming quality (no errors).

User Friendliness

- Is it easy to use? (access t help, return to menu, etc.)
- Is it motivating? (interest for using, appeal, variety of activities, etc.)
- Are the responses flexible? (accepts several, does not require exact matches, etc.)

Finally, and these last questions seem to be the most important, one should ask whether the software impacts positively on performance in the language and whether it enhances students' language skills.

Conclusion

Several strategies seem plausible when implementing technologies in foreign language instruction.

First, using modern authoring systems (including OOP systems) is recommended only when materials are not readily available from publishers. This is also the recommended strategy for less commonly taught languages, for advanced language work and for languages for special uses. Second, the skills taught with the new technologies depend very strongly on the attributes of the technologies. Those attributes should be well analyzed before undertaking a CALL project. Third, advances in second language acquisition theory are beginning to influence CALL materials design. Materials which reflect the new theories are beginning to emerge. Finally, CALL evaluation should factor in not only traditional criteria, but take into account pedagogy.

References

ACTFL proficiency guidelines. (1986). Hastings -on-Hudson, NY: ACTFL Materials Center.

Ariew, R. (1987). Integrating video and CALL in the curriculum. In W. F Smith. (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 41-66). Lincolnwood, IL: National Textbook.

Ariew, R. & Frommer, J. (1987). Interaction in the computer age. In W. M. Rivers (Ed.), *Interactive language teaching (pp.*. 177-193). New York: Cambridge University Press.

Clark, J. L. D. (1988). Toward a research and development strategy for computer assisted language learning. *CALICO Journal, 5*, 5-23.

Goodman, D. (1987). *The complete hypercard handbook*. New York: Bantam Books.

Hope, G, Taylor, H., & Pusack, J. (1984). *Using computers in teaching foreign languages*. New York: Harcourt, Brace, Jovanovich.

Hubbard, P. (1987). Language teaching approaches, the evaluation of CALL software, and design implications. In W. F. Smith (Ed.), *Modern media in foreign Language education: Theory and implementation* (pp. 227-254). Lincolnwood, IL: National Textbook.

Jaffe, C. & Lunch, P. (1989, September). Hypermedia for education in the life sciences. *Academic Computing, 1989*, September. 10-13; 52-57.

Krashen, S. (1985). *The input hypothesis*. New York: Longman.

Jones, C & Fortescue, S. (1987). *Using computers in the language classroom*. Longman.

Lambert, S & Ropiequet, S. (1986). *CD ROM: The new papyrus*. Redmond, WA: Microsoft Press.

Merrill, M. D. (1985). Where is the authoring in authoring systems? *Journal of Computer-Based Instruction, 2*, 90-96.

Schulz, R. (1990). Bridging the gap between teaching and learning: A critical look at foreign language textbooks. In S. Magnan (Ed.), *Issues in Language Program Direction, Vol. 2*. American Association of University Supervisors and Coordinators of Foreign Language Programs. Boston: Heinle & Heinle Publishers.

Smith, W. F. (Ed.). (1989). *Modern technology in foreign language education: Applications and projects*. Lincolnwood, IL: National Textbook.

Woodbury, V. (1988, September). CD-ROM: Potential and practicalities. *CALICO Journal*, 25-33.

Young, R. (1988, March). Computer-assisted language learning conversations: Negotiating an outcome. *CALICO Journal*, 65-83.

Author

ROBERT ARIEW, Associate Professor, Department of French and Italian, University of Arizona, Tucson, AZ 85721. Specializations: computer-assisted instruction and foreign language learning and teaching.

Applied Language Learning
1991, Vol. 2, No. 2, pp. 45-59

Managing Innovation and Change for Instructional Technology

Donald C. Holznagel
Northwest Regional Educational Laboratory

Introduction

Educational organizations are no less likely to be affected by changes in the social and technological aspects of life than any other. The increased rate of change, especially in the development of electronic technologies in the last few years, has been particularly disconcerting to schools, colleges, and training institutions. While teachers are aware that computers and telecommunications tools are available and widely used as a matter of course in businesses and professions, it seems that availability of those tools to teachers and classrooms lags well behind. As new developments occur more rapidly, there is a slightly panicky feeling among education personnel familiar with the technology, that one could even miss an entire generation of computers. While this in itself might not appear to be of much consequence, it can certainly be disconcerting to those who are concerned that education is not keeping up with its social context, and that they are missing the power of modern tools that other professions take for granted. In a concise statement of the position, Lipson (1981) states: "The question is, can we invent new technology-based arrangements that, with different combinations of roles, tasks, reporting lines, schedules, etc. are more productive and more satisfying to both faculty and students?" To be sure, there are many instructors and administrators for whom computers and other technologies are irrelevant or at most peripheral to their central task, and who believe that the standard classroom and standard teacher role are not what needs fixing in education. For some of those, the computer is at most a way of adding interest for students or providing another option for handling drill or other routine tasks within a traditional setting. There is a natural tension, therefore, between two groups of staff: those who believe that change in instructional procedures is needed and desirable and that technology, particularly the computer, can make it possible, and those who believe that changes in the organization and conduct of teaching should be minimal and the computer is at worst a threat or at best a support of the status quo. Most people in an educational organization are probably somewhere in the middle, having both expectations and concerns. An early summary

(Lidtke, 1981) of the reasons that teachers have traditionally failed to use new instructional technologies is interesting to review:

- Little concrete evidence of effectiveness
- Resistance to change
- Lack of training
- Lack of hardware, software, and courseware
- The need to change teaching style
- The need for extra time and preparation
- Anxiety about technical things

Ten years later, after microcomputers are assumed to be commonplace in education, I still find in conducting school district needs assessments that a majority of teachers are nonusers or irregular users of computers in instruction. Furthermore, the reasons given by teachers for the lack of involvement are essentially those described by Lidtke, even including the lack of hardware and software. Although a large number of computers have been purchased in ten years, it is apparently still not enough to have an intensive effect. The instructional technology needs assessment that was conducted by the Institute for Simulation and Training in 1989 for the Defense Language Institute Foreign Language Center (DLIFLC) indicates quite similar results.

Some education agencies, schools, colleges, and training institutions have attempted to keep up with the developments in technology through pilot projects and experiments. Rarely does a school or college have the funding or the will to make sweeping, system-wide changes through the intensive introduction of computers or other technologies. They can usually only afford the time and energy to examine specific aspects or components, so there is little experience with massive change. Furthermore, in instances where pilot projects have been attempted, problems have occurred which might have been avoided by attending to current knowledge of strategic planning procedures and concepts of dealing with change. Much of the literature concerning the change process is written without regard to the specific context of educational technology. However, in applying technology to education and training, a few people have written of their experiences and reflected on the implications for procedures and guidelines. My own experience in assisting school districts and community colleges in strategic planning for educational technology has also provided some notions of the requirements for success, both in the process and the result.

This paper will apply both of those sets of experiences to the introduction of computers and other electronic technologies in instruction. The position taken in this paper is that whether one is interested in changing teaching greatly or in minor modifications to the status quo, evidence is available that computers and other electronic technologies can be valuable. Furthermore, even though new developments are frequent and their implications complex, procedures based on experiences are available which will greatly aid an organization in planning for the introduction of a technology. Such procedures can both take into account a rate of change appropriate to the staff and provide the environment for successful implementation. This paper is primarily about what has been termed "orgware." Gennady M. Dobrov (1979), in an analysis of the organizational context for technology, defines the concept of orgware as the third necessary component of a technological system in addition to hardware and software:

> There is equally a need for a specially planned form of organization corresponding to the level and specific features of the principles and functions realized by the technological system in question under particular conditions of application (what may be called specialized organizational support).

Judging from the needs assessment completed by the Institute for Simulation and Training in 1989, the DLIFLC does not appear to be markedly different from other educational organizations in expressed needs, problems, barriers, and expectations with regard to educational technology. While it may be disconcerting that few others are further along in the process of dealing with these issues, it should be heartening that many are trying and some experience in the implementation of technology has been analyzed by those involved. Organizations even of the same type differ in structure, procedures, personnel, and history so that the process and results of implementing and managing an innovation will vary depending on the goals and measures established.

Approaches to Planning for Change

Two of the questions which were posed in advance to guide the development of this paper provide the focus for this section:

1) What are the essential factors in effectively planning and managing innovation and change in the use of technology for language training?

2) How can participation in planning and implementing technology be effectively broadened and resistance to change lessened?

Review of Advice from the Literature

Several of the references were chosen because they provide different points of view on managing innovation and change. Although none provides a comprehensive plan, each makes a contribution to a larger picture of the aspects of managing the introduction of new technology. A report of the development and implementation of an interactive instructional television system at the Illinois Institute of Technology (Guralnick, 1981) organizes planning issues and questions in three general categories: goals, strategies, and tactics. Within each category, the author lists many questions intended to generate the information required to define the category for an institution. Without listing all the questions, the following paragraphs summarize the content of each category.

Goals: Establishing goals is particularly important since technology is a means to an end, not an end itself. Institutional goals relating to the mission, students and other constituents served, and programs and courses offered should be established. Goals in regard to possible new constituents and services should also be considered.

Strategies: Strategies are long-term (five to ten years) courses of action designed to accomplish one or more goals. Questions in this category include identifying the variety of possible actions which would address a goal, the role of technology in meeting various goals, the barriers which might impede success, and the organizational changes and innovations required to deploy technology. Two questions expressed in this category are really pursuant to examination of goals, and should perhaps be asked in the previous section. The first asks why technology is to be used, and the second asks what internal and external forces are spurring you to deploy educational technology. These are crucial questions to be asked at an early point in the process.

Tactics: Tactics are specific action steps related to one or more strategies. Questions in this category relate to the choice of appropriate technologies, benefits and other impacts, recipients of benefits, training requirements, motivation of potential users, timetable, and funding. Also in this category, Guralnick suggests that in addition to the actual impacts, attention should be paid to who will perceive that benefits and impacts have occurred because of the technology.

The major values of the Guralnick report to the DLIFLC planning process are in its reminders that technology is not itself a goal, that several levels of thought and planning need to be engaged, and that there is a direct relationship between success in implementation of

technological innovation and the amount of attention paid to the three categories and their interrelationships.

A research report (Watkins, 1982) analyzes the elements in a project approach to introducing innovation. Specific managerial personnel positions having distinct responsibilities critical to project success are as follows:

Project Director: Trains users and solve operational problems.

Organizational Sponsor: Provides support and organizational vision.

Watkins found that the "figurehead" role of an organizational sponsor is an important aspect of leadership. Watkins also identifies three stages of managerial concern in a project as initiation, implementation, and institutionalization. The study attempted to determine the managerial roles in the stages of concern which were related to successful innovation. Some very practical advice for managing change resulted from the study, including:

- Managers should be selected for their interpersonal skills.
- Managers with skill in innovation should govern a project at implementation, while a manager with skill in maintaining a program should govern a project at institutionalization.
- Managers should be provided training in the nature of innovation and the implications for support.
- When innovation is nonroutine, a project should be placed under the sponsorship of one who is able to invent creative entrepreneurial solutions to problems.
- With nonroutine innovations, strategies from management to reduce stress may be helpful.

An analysis of management concerns based on conducting extensive research, development, and implementation of computer-based instruction in a large school district (Estes, 1981) resulted in a categorization of tasks from an organizational rather than project perspective. Estes was a superintendent of the Dallas Independent School District in Texas, stimulating and managing the extensive development of instructional computing activities over many years. He separates managerial concerns and problems into six areas.

Political: . He recommends paying attention to the political climate for technological change as it applies to those people or agencies providing the funds and to those staff members who must carry out the implementation. With regard to teachers, he believes that assurances

need to be given that technology would complement rather than replace them.

Fiscal: Funds must be sufficient to support an effort calculated to make a difference. New technology costs considered as an add-on to current costs are still high even though the capabilities of computers are increasing while costs are coming down. If teachers are not to be replaced, people might expect that at least productivity gains will be realized from the extra expenditure.

Technical: In this category, Estes emphasizes the importance of having qualified technical staff who can interact well with educators to support the staff. Cautions are that new developments should be restricted to pilot test settings until debugged and evaluated, and that one should avoid buying any new technology just to be in the "in" group.

Instructional: Of prime importance in this category is that the curriculum framework must be developed before entering into the development of technology applications. Curriculum developers and supervisors need to be thoroughly trained regarding the potential and limitations of the technology. The ultimate users—teachers and students—need to be brought in at an early stage in the development and testing process for orientation to the new technology and feedback to the developers.

Policy: Governing board and administrative policymakers must be brought along in their understanding of the value of the new technologies through demonstration and hands-on experience where practical. Board policies should focus on what needs to be done, and administrative policies on how it is to be done. Finally, board policies should be updated periodically to encompass new technologies or major changes in technology which may affect the plans for implementation.

Management: In the school district, Estes established the position that the management of technological change was one of the primary missions of administration. He recommends that administration at all levels should demonstrate reliance on and confidence in new technologies, and that administrators must become directly involved with teachers as technological changes are implemented. A caution is raised that new management processes may be required in order to effectively monitor and assess the effects of the changes.

The categorization of management concerns offered in the Estes paper constitutes a very useful way of planning administrative responsibilities in any educational organization in the development of a plan for implementing technology in instruction. Another analysis (Elsroad, 1981) has been written by a superintendent of the Montgomery County, Maryland School District, which has a 25-year history of

extensive research, development, and implementation of computers in instruction. There is understandably much overlap between Elsroad's experience and that of Estes, although it is presented in different terms. He offers a set of very practical recommendations. A selected subset is paraphrased below.

- Efforts at change should grow out of well-defined real problems in order to gain the support necessary to sustain them.
- Objectives should be both long- and short-range, simply stated, understandable, meaningful, and achievable.
- Prepare a comprehensive plan for building understanding of proposed change and have a broad base of participation in the preparation of the plan.
- Project cost realistically and report regularly to avoid distrust. Cost effectiveness is increasingly important in education.
- Provide the management support system necessary to maintain the program.
- Training requires special attention, and must help people cope with the personal behavior changes they face.
- Staff the program with a leader knowledgeable in both education and technology who is also a good manager.
- Report and discuss progress several times per year, and relate progress to objectives in the master plan.
- Progress should move forward boldly with constant momentum, but not more than can be well-staffed and supported.

In addition to the work of the previous authors providing specific advice on plans, procedures, and people in the change process, it is helpful to have a philosophical frame of reference within which to work. While it is good to have specific plans, it is difficult to develop them because of the characteristics of increasingly rapid change in the current social and technological environment. Because both the new technology and the instructional process are complex, an implementation plan may cover ten years, while at the same time significant developments in electronic technology can occur every year. A report of a Teacher Corps project involving educational change (Pine, 1981) discusses the selection of a management philosophy called Systematic Ad Hocism to guide the implementation of research-based innovations. The situation described

in the project was similar to that for the introduction of a new technology. It was necessary to develop a plan in order to accomplish a goal, while simultaneously providing the flexibility to deal with new developments and the resolution of uncertainties during the execution of the plan.

The report identifies four major characteristics of Systematic Ad Hocism as follows:

1) *Having a map rather than an itinerary*

Maps are general conceptualizations, comprehensive in coverage of the organization and the changes sought, while itineraries are prescriptions for management or planning. Itineraries limit freedom and cannot handle the unexpected events in a complex, changing environment. A map allows for the flexibility of alternate routes to a destination if unexpected events intervene in the one first chosen.

2) *Being long-range*

Long range planning includes forecasting and generating alternate futures, which provides a context for devising short range solutions to problems as they emerge over time.

3) *Being adaptive*

Adaptive planning has two aspects. First, it is being adaptive to the environment of the change, working with the implications of the change. Second, it is an active response to challenge the environment, making choices to meet original goals as experience with the change grows.

4) *Being value-based*

Participants in a change process are assumed to be not neutral, having their own ideas about instruction, the organization, their subject, and change. It enables staff members to have confidence in their ability to be part of a team by recognizing their ability to contribute to the plan as well as the execution.

The philosophy is gradualist, seeking to change without explosive disruption, and typifies the historical experience of change in education. The project leadership purposely selected the philosophy because of the nature of the project, and found that it resulted in some success and some failure. In the report, Pine used the successes and failures to examine the nature of change ideologies. For the purposes of dealing with technological change in an instructional institution, the philosophy supports an approach which plays down the threat to teachers by allowing for their participation in the processes of planning, experimentation, evaluation, and implementation, and by recognizing the value of their talents and abilities.

Author's Experience

The foregoing summaries collectively represent good advice from several points of view. The following points are based on my experience in school district and community college strategic planning activities. They represent my perception of the priority areas of concern in planning for technology. They incorporate some of the advice from previous references. However, just because many items from above are not repeated here does not mean they are not important to keep in mind. I am recasting the advice and adding my own perceptions to address the planning situation as I have found it in educational institutions, particularly to address deficiencies I have observed which retarded progress in those institutions.

1) *Have A Plan for a Plan*

An institution needs to have a planning approach. I believe that the three steps suggested by Guralnick of setting goals and objectives, developing strategies, and developing tactics constitute a simple but effective framework for planning. If one is left out, the others are useless exercises. If you establish goals but have no strategies, action is impossible and staff are frustrated. Strategies and tactics without goals may result in actions which are likely unmeasurable and without long-term effect.

A basis for the planning approach may also be supported by a general philosophy or set of principles, as in Systematic Ad Hocism, discussed above. At the very least, it needs to be clear whether drastic change or gradualism is the chosen path, and if the latter, how gradual the path is to be. I believe that in the implementation of computer-based technologies, the threat of teacher replacement is frequently such a strong factor that gradualism of some sort is required. This does not mean change cannot occur in a reasonable time, however. Rather, it means that a staff training and pilot demonstration program must be as robust and successful as the proposed rate of change is ambitious.

2) *Set Realistic Goals*

Goals are not achievable in any plan if they are seen to be irrelevant or unbelievable to the people who must try to reach them. Thus, if we set forth a map for change and the map is on too large a scale, it carries insufficient information and guidance for people in the institution to make their way. A map can be on a suitably small scale and still offer many options for paths or itineraries. In the case of the DLIFLC, a goal such as "improving foreign language instruction" may be desirable but does not communicate achievability. It is probably a nice place to go, but no one can devise how to get there. On the other

hand, statements like "improving the efficiency of language instruction in the DLIFLC" or "improving spoken language skills in students of the DLIFLC" can stimulate immediate concern about how, why, and when they can be achieved, and lead to a discussion of strategies and tactics. In the goal-setting process, it is important at a very early stage to determine the basis for interest in the implementation of a technology. Why is the institution moving to employ computers, video discs, or CD-ROM? Is it because of teacher interest, or demonstrated results in the professional community, or because their capabilities match with needs such as efficiency? I am not arguing in favor of a particular reason, but rather that the process of identifying the reasons will help in establishing realistic goals and will build the rationale necessary to enlist the support of all affected groups, including students, teachers, and administrators. The participation of all those groups in the process will invest the goals with even greater legitimacy.

3) *Plan for Change*

The plan for technology should include a plan for changing the plan. This means building in the procedure for evaluation, review, and modification. It should include an examination of all aspects of the activity as described in the Estes paper (1981): political, fiscal, technical, instructional, policy, and management. Such review should be at least annual, but depending on the scope of the activity, could be more frequent. It is likely that some components, such as hardware performance or teacher training for example, should be reviewed more often than others. The administrative leadership should be prepared to increase the review frequency if problem frequency increases or information is insufficient. Including this component does not presuppose that there must be changes to the plan or that the whole plan will be changed every year. However, it is reasonable to expect, as pointed out before, that both technology and people will change, and sometimes very quickly. Computers sometimes break, and people sometimes burn out or change jobs. Those two concerns alone are sufficient to establish regular, frequent, mandated review points in the plan.

4) *Develop a Modular Implementation*

As mentioned earlier in the paper, rarely does an agency have the luxury of institution-wide implementation of new technology at one time. However, most people want some measurable effect out of any large expenditure. A major problem occurs when the technology is seen as somehow separate from the real work of the institution, and is treated as an end in itself. I am referring to the approach to implementation which tries to base purchases on a notion of equity, such as buying a set number of computers per department, or identifying an amount of

money to get spent equally across departments. Since such a plan has no relation to goals or to an analysis of requirements, then no result can be measured and related to goal achievement. It is far better to spend available resources on meeting one goal or on fully implementing a few strategies, because progress will be shown and success can likely be demonstrated. Therefore, if you cannot fund and carry out an entire plan at once, I recommend that you develop the plan in modules, implementing one or a few modules fully each year over three to five years. This will have additional benefits as noted in the next item.

5) *Establish Meaningful Cycles*

Enough is known about electronic technologies and the DLIFLC organization that the plans for implementation and review can be designed to accommodate patterns and cycles which are based on other realities. Based on history for example, computers can be expected to have a useful life of five to seven years. Within that time, the machines may wear out, or incur high maintenance rates, or become unsupported by the manufacturer, or have insufficient capacity to handle new software. Those are predictable happenings.

Suppose it is also the case that the curriculum for German language instruction including instructional materials is reviewed every five years, a typical cycle in school systems. Since technology is not an end in itself but an instructional tool, it makes sense to consider computer hardware and software in the curriculum review. Software probably has a longer useful life than hardware, but it is unpredictable. It does not wear out in the hardware sense, but it can be made almost useless because of changes in the philosophy or pedagogy of language instruction, occasioned by staff or textbook changes or other forces. In any case, if the computer is demonstrated to be a valuable tool in a curriculum and is accepted and used as an integral part of the instructional process, then it should be considered and reviewed in the same manner as texts, films, or other tools. Amortization of the cost over the chosen period and an expectation of replacement or upgrade costs are integral to such consideration.

A five- to seven-year cycle may also make sense for staff development and readiness reasons because that period of time has been suggested as necessary to become comfortable with using computers. Dr. Karen Sheingold, Director of the Center for Technology in Education at Bank Street College, reported in a recent conference presentation that a survey of teachers having lengthy experience with computers in instruction indicates that about five to six years were required for them to become comfortable with integrating computer applications with instruction. Any introduction of a new technology to a group previously unfamiliar with it needs to take that experience into account. Realistic cycles will also provide information for decisions on pilot activities and

implementation. Included in an overall plan might be a staggered implementation, so that the financial impact of investment in technology and related training can be spread over a period of time, and the similar impact of replacement can be anticipated in the plan. An additional benefit can accrue from such cycles, in that the technology is seen to be a part of regular consideration in a logical framework. The pressure to replace equipment with the latest thing can abate because staff are assured that "their turn" will come in a reasoned fashion and they can focus at that time on the best available version of the technology at that date. It also places the responsibility on the planning staff to make the best possible recommendation at the time because they will be expected to use it through the cycle. It should be noted that a mid-cycle review is useful to include in the plan because the notion of a useful life for a device is expressed in averages. When we say that five years is a useful life span for a computer, that means that probably some failures in some components could appear in the third year. Hence, a three-year review point to check on the rate of failure is prudent.

Other Issues and Considerations

Several questions in addition to the two addressed in section II were posed in relation to the general topic of managing innovation in technology in the DLIFLC. They are more specific than the others in addressing particular aspects of implementation, and will be dealt with individually here.

1) *What is the best balance between R&D and language training technology program implementation efforts?*

There is no single best guideline for R&D investment regarding instructional technology. That is, there is no standard percentage of overall investment for R&D which organizations use to assure success. Rather, what is "best" for a given situation is related to what is required to carry out a strategy and achieve a goal. R&D activities generate knowledge, and implementation activities make use of that knowledge. Hence, in a given situation, the question must be asked at the start: "If we decide on this strategy, is there sufficient knowledge of its effects to carry it out, or are there gaps in the knowledge base necessary to a full implementation?" To use a previous example, suppose a goal has been set to increase the efficiency of language instruction, and further suppose that part of the strategy is to use computers to increase the vocabulary practice on an individual-student basis. If we ask whether software is available to carry out the task, and the answer is either "no" or "an insufficient amount," then a development effort is in order, followed by a research task to determine if the development is effective. If the answer is "yes, but there is no evidence to say it is effective," then

a research task is also appropriate. It is likely that both cases will occur in the field of foreign language instruction. It seems reasonable to assume that because the DLIFLC is responsible to teach a number of less commonly taught languages, there is probably not a ready supply of proven courseware available for those languages. It is also possible that where courseware is available, the instructional approach may not match the approach preferred by teachers in the DLIFLC, or it may be at a level too elementary for DLIFLC needs. Hence, it can be assumed that an ongoing R&D effort is warranted, but the amount of such effort in any given year may vary with the stage of implementation of the various modules in the overall plan. The existence of a plan should provide the information to forecast the R&D requirements.

2) *How should a training institution select, obtain, and place hardware and software to serve the functions of instruction, staff development, and R&D?*

The selection, acquisition, and placement of hardware should be directly related to the notions of modular implementation and meaningful cycles addressed in section II. Technology, especially computers, should be acquired in amounts dictated by the requirements for success in implementing a module. Buying all the devices required for the same purpose in the same configuration at the same time pays dividends in standardizing software, instructional materials, teacher training, and ongoing support and maintenance. If that pattern is repeated for each module, then even if there are differences in the devices between modules, there will be consistency for everyone concerned with the same implementation activity in the same year. A major concern in hardware purchase is obsolescence. It cannot be avoided, but planning can minimize its effects on the instructional program. The concept of expected useful life was discussed in section II. It can allow the institution to plan replacement. However, it is likely that at the end of the expected life, a number of working units still remain. In the case of computers for example, it is likely that of a given collection, some will still be working at the end of seven years. If so, a strategy for effective use of that equipment can be devised, using the concept of consolidation and repurposing. As an example, suppose that a group of 20 microcomputers are in a lab used primarily for vocabulary instruction, and that after six years, 12 of them are still working reliably. There is no reason to stop using them if the same function or another function for which software is available needs to be addressed, but it may be the case that those same computers are not marketed and none can be bought to rebuild the collection. The regular review process should identify the replacement for the collection and the remaining working machines should be placed in such a way as to be useful for their original or a newly defined purpose as long as they are working.

The placement of hardware should also be governed by the goals and strategy. There is no research to say whether computers should be placed in labs or in the classroom. However, there is a case to be made for any configuration based on a clear idea of instructional approach. A single computer in the classroom is effective for large group demonstration, teacher records, and materials generation. A computer lab is useful for large group activities where each student needs a station, and to handle a large number of students having a variety of individual uses. Special configurations of computers may be needed to support database searches, telecommunications, or other applications. Analysis of the instructional strategies will provide the information required for effective placement.

3) *What are the resource requirements for effective technology implementation? How can these resources be justified and obtained?*

The resource requirements are determined by the institutional plan. Effective technology implementation means that there is an expected effect of technology in relation to goals and strategies. If the plan identifies the desired goals and strategies, the amount and type of technology must also be identified. Further analysis can identify the cost of hardware, software, personnel, training, facilities and other implications.

References

Bennis, W. G., Benne, K. D., & Chin, R. (1985). *The planning of change* (4th ed.). New York: CBS College Publishing.

Dobrov, G. M. (1979). Technology as a form of organization, *International Social Science Journal, 21*, 585-605.

Elsroad, H. (1981). Managing technological change in Montgomery County, Maryland, In *Technology and Education: Policy, Implementation and Evaluation: Proceedings of the National Conference on Technology and Education* held in Washington, D.C. by the Institute for Educational Leadership.

Estes, N. (1981). Managing technological change in the Dallas public schools. In *Technology and Education: Policy, Implementation and Evaluation: Proceedings of the National Conference on Technology and Education* held in Washington, D.C. by the Institute for Educational Leadership.

Forsythe, K., & Collins, V. (1983). *British Columbia—higher education*

and the integration of a new technology: Case study. Paris, France: Organization for Economic Cooperation and Development.

Guralnick, S. A. (1981). Deploying educational technology at an independent urban institution. In *Technology and Education: Policy, Implementation and Evaluation: Proceedings of the National Conference on Technology and Education* held in Washington, D.C. by the Institute for Educational Leadership.

Lidtke, D. K. (1981). Securing teacher acceptance of technology. In *Technology and Education: Policy, Implementation and Evaluation: Proceedings of the National Conference on Technology and Education* held in Washington, D.C. by the Institute for Educational Leadership.

Lipson, J. I. (1981). Changing organizational structures to capitalize on technology. In *Technology and Education: Policy, Implementation and Evaluation: Proceedings of the National Conference on Technology and Education* held in Washington, D.C. by the Institute for Educational Leadership.

Pine, G. J. (1981). *The certainty of change theory: An analysis of experience and ideology*. Paper presented at the Annual Conference of the Association of Teacher Educators, Dallas, Texas. University of Central Florida, Institute for Simulation and Training.

Pine, G.J. (1989). *Resident training: Instructional technology needs assessment, prepared for the Defense Language Institute Foreign Language Center*. Orlando: University of Central Florida.

Watkins, K. (1982). *Managing change: How managers describe their roles and concerns*. Paper presented to the Institute for Presidents, Austin, Texas.

Author

DONALD C. HOLZNAGEL, Director of the Technology Program, Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, OR 97204. Specializations: curriculum, educational computer applications, courseware evaluation.

Applied Language Learning
1991, Vol. 2, No. 2, pp. 61-75

Software for Language Training: Directions and Opportunities

James P. Pusack
The University of Iowa

Introduction

The language teaching profession knows far more about bad software than it knows about good software. Over the years, in fact, the profession has generated a comfortable pool of wisdom about computer-assisted language learning (CALL). We think we know that flashcard-like drills are beneath contempt; we chisel stone tablets proclaiming that software that simply replicates older technologies (i.e., book and blackboard) cannot have value; computerized Hangperson ranks at the bottom of the technological food chain; games that shoot off rockets to reward the correct adjective ending embarrass us; and hour-long adventure games that expose the learner to minuscule amounts of contextualized practice appear manifestly inefficient.

These miscellaneous examples of our putative knowledge, however, also reflect the extremely tenuous nature of previous collaboration between specialists in language acquisition and foreign language software developers. Are there no drill formats based on sound psychological principles that might prove more efficient than learners' own haphazard attempts to master vocabulary? Are there no reasons to replicate excellent teachers with excellent courseware? Is there no room for game-like environments that focus students' attention on the task, rather than on words and structures? Should young language learners who like to play Hangperson in English be forbidden from playing the same game in a second language?

These few examples already demonstrate the dearth of any absolute principles to guide the development of language instruction software. Keeping this caveat in mind, one should be able to discover some non-absolutes based on current theories and current technologies. The following section will attempt to enumerate some essential characteristics of effective language instruction software based on theoretical models, instructional efficiency, and programmatic needs and resources. A separate section addresses special issues related to strategies for accessing multimedia and mass storage devices. The next

section deals with the role of productivity tools and authoring systems and is followed by a discussion of strategies for selecting, developing, and adapting software. The thorny but critical problem—software/hardware compatibility—is treated briefly. Finally, the concluding section offers three recommendations for administrators who hope to establish directions and exploit opportunities for foreign language instructional technology. In many of these sections, an attempt has been made to compress complex issues into manageable form by compiling lists of concise recommendations and observations in quasi-aphoristic form. It is hoped that such brevity will generate both heat and light.

Essential Characteristics of Effective Language Instruction Software

One shibboleth we frequently encounter states that a language teacher who can be replaced by a computer should be. Or is it that a computer than can be replaced by a human should be? Whichever turns out to be correct, three complementary and overarching principles remain to guide us in planning and developing software for computer-assisted language learning:

- There exist general principles of language teaching/learning/acquisition that apply equally well to both people-mediated instruction and machine-mediated instruction.
- Some language-learning tasks are best performed with human instructors, some with computers, and some with both.
- People, not machines, will determine what learning environment works best for a given set of goals and learners, so that in the end all instruction, including CALL, is people-mediated.

The remainder of this section describes various characteristics of good CALL under three headings derived from the above three principles:

1) theory-related characteristics; 2) efficiency-related characteristics; and 3), program- or site-related characteristics.

Theory-Related Characteristics

How can we define the essential characteristics of good or effective software in a theoretical and methodological vacuum? We cannot. Worse yet, even in a language-learning setting where basic principles and objectives have been established, the essential characteristics of newer instructional technologies are so little understood today that it is a challenging task to draw lines from teaching/learning points A, B, and C to computer-interaction types 1, 2,

or 3. For example, if a skill priority has been set for the ability to scan text for specific information or main ideas, it is easy to reject any approach using traditional multiple-choice items with a stem and distractors, but it remains much less obvious what approach might be more appropriate, not to mention what software design might be ideal.

The only solution to this conundrum, therefore, lies in postulating that the essential characteristics of effective language instruction software are, at some level, the same as the characteristics of good teaching, effective learning, and successful acquisition. To put it more provocatively, good software should constitute an ideal learning environment. It should do what a teacher and/or a learner would like to do under ideal conditions, as established by the prevailing language learning model. This aspiration echoes the general and too-familiar claim that computers individualize instruction by providing immediate feedback. The intent here is to challenge the experts on language teaching: Foreign language software designers must fling down the gauntlet at the feet of methodologists and demand rigorous definitions of those interactions derived from theory that will lead to language acquisition. A conscientious designer of CALL for listening comprehension might cry out: "Don't keep telling us to administer massive doses of comprehensible input! Tell us what size dosage of what type activity will move a student from 1 to 1+1! And be sure the hypodermic you want us to use doesn't raise the patient's affective filter."

In the absence of any universals fully acceptable to the whole profession, and lacking the requisite specificity about methods derived from such universals, CALL designers find themselves in the position of drawing tentative dotted lines between, on the one hand, what appear to be useful approaches and, on the other, the kinds of interactions the hardware and software can provide under the real-world constraints of time and money. Currently productive research directions that can lead to long-term progress in CALL include:

1) efforts to refine national proficiency standards and extend the broad proficiency concept to all four skills;

2) insights arising from language acquisition theory and research;

3) strategies based on schema theory; and

4) recommendations concerning effective learner strategies. Under these conditions, we can expect progress to be made by those CALL developers who attempt to do one or more of the following:

   - provide communicative activities
   - contextualize practice

- subordinate grammatical correctness to successful expression or comprehension;
- impart cultural content as well as skills;
- expose students to authentic language and texts;
- take into account the differences among learners; and
- build or reinforce transferable learner strategies.

Given the profession's notorious and oftentimes unproductive commitment to disagreement about major principles, nothing could be more foolhardy than to ignore the fundamental need for sound theory and compile an "interim" list of guidelines for effective software. Such a list follows:

a. Many good interaction types in foreign language computing can be transferred across languages; in the European languages, virtually everything should be transferable.
b. The best software should present language in a contextualized, discourse-level format, preferably based on authentic print and video texts.
c. At the lower levels of language learning, the apparent conflict between maintaining motivation (low affective filter) and providing a high level of authenticity can be resolved via computer-direct tasks and helps.
d. Software that accounts for differences in student learning styles, motivational levels, and preferences has the greatest chance of succeeding.
e. Software that downplays explicit discussion of grammar principles will be adaptable to a wide range of settings and to changing approaches and textbooks.
f. Software, especially that for the receptive skills, should allow choices that reflect student interests.
g. Hypermedia-style courseware may provide interesting ways to explore cultural artifacts, but may not be effective or efficient in building language skills; there are no easy answers to the many outstanding questions about the issue of learner control versus program control in CALL.
h. Comprehensive tutorial software that could handle all, or even significant portions of novice/intermediate instruction has not yet emerged and may not be a prime goal.
i. Acquisition will take place most readily when activities are content-based, rather than focusing on overt skill-building; technology can structure a smooth transition

between learning and acquisition of listening and speaking skills.

j. Good software offers a high level of remediation, feedback, and diagnostic information available to either student or teacher.

k. Good software should be designed to foster a feeling of success in the user.

l. Primitive forms of remediation can raise affective barriers.

m. The place of instructional technology along the continuum between learning and acquisition remains to be defined by significant software packages.

n. In exercise-type software, psychological principles related to short-term vs. long-term memory dictate management techniques such as recycling of items or item types.

o. The role of artificial intelligence (AI) in CALL (particularly expert systems and natural language processing) remains to be demonstrated.

p. Courseware that uses judicious combinations of input via the keyboard and pointing devices and output via text, graphics, multiple windows, color, recorded sound, and video (multimedia) promises to provide highly effective solutions. Significant (non-cosmetic) uses of voice input for CALL depends on advances in AI; synthesized speech (text-to-speech) at the sentence level has similar problems.

q. Not all potential uses of technology for CALL are equally suitable for all levels of language learning.

r. Software that allows detailed tracking of learner behaviors (not just performance levels) will foster a better development process and lead to significant refinements in the software itself, as well as providing a sound basis for second language acquisition research; this shall hereafter be called the *Garrett Imperative*.

s. Software that explicitly encourages and structures small-group work will tend to build multiple skills simultaneously and motivate communicative processes.

Efficiency-Related Characteristics

These characteristics derive not primarily from attempts to reflect sound (i.e., research-based) methodology, but rather from considerations about the best match among various technologies. Even when it is technologically feasible to have computers offer some learning activities, in most situations there are general principles that dictate

needs and priorities. Such judgments will change over time as the demonstrable capabilities of CALL improve and expand, so that, for example, the tradeoff between class/teacher time and lab/homework time eventually balance out on the side of computer-based activities in a laboratory or home setting.

A clear example of this phenomenon is the use of computers to evaluate student essays. While major projects are underway to create various pedagogical parsers that students could use as tools in preparing foreign-language compositions, the scarcity of such tools, their limited usefulness, and the cost of student workstations needed to run them preclude widespread implementation of computers in this aspect of composition instruction: Teachers must still grade compositions. In five years, however, these factors will change to the extent that teachers will no longer need to correct errors in most low-level grammatical structures; instead, they will evaluate such higher-level aspects of student composition as lexical choice, idiomatic usage, style, rhetorical features, and—above all—content.

In order to explore what uses of instructional technology will generally be most effective, efficient, and cost-effective, it is useful to juxtapose two dichotomies—the learning/acquisition dichotomy and the productive/receptive skills dichotomy. For many language acquisition specialists, both these pairs represent complex, interrelated aspects of language study. Yet in examining just where computers have the most promise, the distinctions can prove helpful. We need to determine just where in the following chart instructional technology can best provide solutions that improve on current approaches:

FIGURE 1

Efficiency-Related Characteristics

SKILLS	Learning	Acquisition
Productive	I	II
Receptive	III	IV

Traditional drill-and-practice of grammatical rules as a means of building the speaking skill falls into quadrant I; this is where much of current CALL falls and where enhanced CALL using artificial intelligence techniques to evaluate student writing, as described above, will eventually be located. It is important to note that no particular quadrant contains the "good" CALL while another contains the "bad" or ineffective. What can be discerned, however, is the area where CALL will have the greatest difficulty: quadrant II.

Quadrant II contains the active, spontaneous use of language for expression. Despite the best efforts to devise complex simulations involving multimedia, authentic language, speech recognition, speech synthesis, and the like, it does not seem plausible that instructional technology will ever fully enter quadrant II for the speaking skill. The writing skill will also suffer from similar constraints unless fundamental advances can be achieved in natural language understanding; in other words, even the most exemplary work on the productive skill will remain in quadrant I—the building of sub-skills and learning strategies for speaking and writing. To be sure, some applications of technology to the productive skills will allow for students to engage in free expression. Multimedia, for example, can serve up excellent forms of simulation for conversation. Databases of lexical and grammatical information can provide useful tools for writing. Nevertheless, most of the best interactions we can imagine for the productive skills will not leave quadrant I far behind. To put the harshest turn on this analysis: It is not clear that simulations will ever do anything for language study but simulate. It is not yet established that simulations promote acquisition.

The simple conclusion from this argument is that live instructional settings will always be most viable and cost-effective as acquisition environments for the productive skills. Teachers and students will always be the best hearers and readers of one anothers speaking and writing. To the extent that real production must be communication, it requires human response. Any technology that offers less than this will be ineffective. The appropriate role, therefore, for technology in the productive quadrants is in building fundamental skills (learning) and in providing tools for teachers and students to do good writing and speaking.

If the human instructor predominates over technology in quadrants I and II, the roles are potentially reversed in quadrants III and IV, which cover the receptive skills. Where the teacher, even in a one-on-one setting, has extreme difficulty in managing the listening or reading process of the student, computer technology can not only provide listening and reading materials, but also structure the skill-getting and skill-using processes in radically new ways. In a computer-based receptive-skills curriculum, learning and acquisition are in constant

flux. Tutorial-style environments in which the technology acts as a surrogate instructor will tend to fall into quadrant III, while hypermedia-like explorations under high learner control will fall into quadrant IV. At the extreme right of quadrant IV we might place truly recreational reading or extensive reading of long fiction where the use of computer-based reference tools has been phased out in the interest of motivation and practicality. The receptive aspects of conversation, finally, fluctuate between the upper and lower quadrants. The role of technology in conversation has aspects of all four quadrants.

At this stage in the development of CALL for the receptive skills, the similarities between software for reading and software for listening are more striking than the differences. Interactive video and audio can deliver listening materials with nearly as much control and flexibility as high-resolution screens can present text. Aids to listening and reading may be quite similar. Tasks and desired skills may overlap greatly. Finally, multimedia settings encourage the creative mutual reinforcement of the two skills, with video providing a visual channel of information to establish and reinforce meaning.

To the extent that sophisticated instructional technology can provide highly effective courseware in quadrants II and IV (with the exception of conversation), one could speculate that the most appropriate use of human expertise lies not in functioning as some sort of eventual ancillary to the machines, but in designing, creating, developing, adapting, and evaluating high-quality reading and listening materials. The ideal instructor's day might therefore consist of 40 percent in developing materials for quadrants III and IV, and 20 percent in compensating for the weak points of the prevailing technology in all quadrants via teaching and/or materials development.

Program-Related Characteristics

People, not machines, will determine what learning environment works best for a given set of goals and learners, so that in the end all instruction, including CALL, is people-mediated. This rather homiletic statement is meant to call attention to the fact that neither theory nor efficiency alone can dictate the best forms of CALL for a given institution. What is missing from the complete formula is a grasp of the needs and resources at hand. Each language program has peculiar objectives, unique advantages, and limited resources. A university with widely disparate students, strong experience in software development, faculty expertise in computer science, language acquisition and software design, and no particular mandate to solve any specific problems via computer, may choose to embark on a long-range project to create a comprehensive, multifaceted foreign language instructional system that can manage many skills and goals and can adapt to changing national

priorities. An adult education course that promises to teach survival or tourist German via one-night-per-week meetings in eight weeks will probably not use any technology at all.

Given this span of possibilities, generalizations about site-specific issues are as risky as those that attempt to transcend theory. Here is a list of observations and recommendations:

a. In the non-immersion residential setting, some technologies (such as video, interactive video, and satellite) can effectively complement in-class exposure to authentic language and help foster an immersion-like environment.
b. For optimum student motivation, software must be close to the curriculum.
c. Resources must be dedicated not only to the development or acquisition of materials, but also to discussion of how technology should be integrated into the curriculum.
d. The match between the curriculum and the software is probably more important than the absolute quality of the software; by waiting for leading-edge applications one can sacrifice opportunities to teach better and thus to gain useful experience with technology.
e. Good CALL tasks and interactions have a high degree of return for the investment of student time.
f. Student tolerance for highly focused language-learning tasks varies greatly, but is usually far higher than faculty tolerance for such tasks.
g. The more closely a piece of software relates to immediate learning goals and real-life tasks, the more likely it is to generate heavy student use.
h. Not all uses of technology need to be fully self-contained; instructors can assign sophisticated written or oral tasks which require interaction with video and computer materials but which are performed in small groups, in individual written work, or in traditional class interactions with the instructor.
i. Not all multimedia approaches require interactive video; students can be asked to watch a videotape at home and then perform computer exercises based on the context provided by that video; if computers are more available than video devices, instructors can show the video in class and require that computer assignments be completed outside of class.
j. Some uses of technology are justified simply because

they provide an alternative delivery mechanism to the one the students are accustomed to in the classroom.

k. The availability of equipment, student work habits, and teacher qualifications to develop instructional materials will all force compromises in any plan to implement CALL on a wide scale.

Considerations such as these will provide a realistic perspective on theory-based aspirations to revolutionize language teaching via technology.

Software Strategies for Accessing Multimedia and Mass Storage Devices

We are only beginning to see persuasive examples of multimedia and CD-ROM. These examples suggest that extensive use of instructional technologies for building the receptive skills will not be hampered by any technical constraints. We are now able to call up and manipulate vast quantities of sound, images, and text in ways that fully meet our needs. The problems of these new technologies are not so much technical as administrative and creative.

A major benefit of multimedia derives from the ability to retrieve the right data at the right time. Whether we use a videodisc to control the replay of authentic television broadcasts or use a CD-ROM to retrieve texts, reference materials, or digitized sound, a significant problem arises from the need to solve copyright problems for the archived material. On a small scale these problems can be solved by creating one's own materials and devising appropriate interactions; on the large scale that justifies investment in these media, however, the rights holders of the materials must be recognized and remunerated.

Strategies for multimedia and mass storage that do not involve retrieval of pre-existing information dictate an extremely heavy investment in materials development. Often the creation of images, video, and other databases lies well beyond the expertise and resources of any single institution.

In either case (obtaining rights or creating new materials), the cost of most multimedia projects suggests collaboration across several institutions. Even when sufficient funds from a single institution can be dedicated to a large multimedia project, the complexity of the effort demands shared expertise, extensive collaboration in the design, and wide testing so that the idiosyncracies of a single institution do not restrict the usefulness and longevity of the resulting product.

The mandate to conduct multimedia projects from an inter-institutional perspective suggests that the development platform used for such projects should not be chosen with an eye only to the technical possibilities. Hardware bells and whistles can make a project more

exciting, but can also isolate the project from the small community of collaborative colleagues who could contribute to the quality of software in development.

Mass data storage raises issues of interaction style. One popular model for the use of large databases of text and images is the hypermedia mode, which provides an exploratory environment that allows students to explore connections among loosely related assemblages of materials. In a university-level setting which expects students to engage in various levels of research beginning with freshman rhetoric courses and extending through completion of the major, hypermedia can foster creativity and stimulate the search for knowledge for its own sake. In a more highly directive training environment it is not yet clear that extensive learner control will prove to be a cost-effective means of building language skills, although aspects of the hypermedia model could be adapted to form an open-ended, browsing environment for acquisition of the receptive skills.

Access to high-capacity storage devices also raises the issue of networking. In a residential setting it is natural to assume that students will have workstations in their own living quarters and that these stations will be able to access text and graphics via networks. A language-learner network could also link students to instructors for tutorial purposes and to other students for purposes of peer exchange in the target language, which could act as a naturalistic acquisition environment for reading and writing.

The Role of Productivity Tools and Authoring Systems

The current status of productivity tools for CALL is not satisfying. Many development aids stifle, rather than stimulate, creativity. To be sure, each generation of computers brings with it new breakthroughs. We can now interact quite effectively with text-like objects on a screen, use pointing devices to move words and focus on sentences, even animate syntactic and lexical relationships using only authoring systems or scripting tools. Control of videodisc players and sound resources has become relatively straightforward. Nevertheless, the complexity of the most interesting tasks we hope to see on our current generation of courseware taxes the capacity of authoring systems and languages to the point where both development time and interaction time can frustrate the user.

As a result, general-purpose productivity tools appear to be most valuable as prototyping mechanisms. The more complex the set of media we hope to integrate into a single, intuitive learning environment, the more crucial it is that non-professional developers (i.e., faculty) be able to experiment with the look and feel of possible interaction types. Prototyping tools make this possible, but will often be

employed as an intermediate step in the development process between design and programming.

Because of the high cost of software development and the limitations of general-purpose authoring tools, many institutions have chosen to develop dedicated foreign language authoring systems or templates. All such systems are far more constraining than general-purpose tools, but they are extremely cost-effective as a way of creating large amounts of high-quality, reliable courseware. To the extent that faculty can accept and adopt the kinds of interaction offered by a particular CALL authoring system, they will forgive or overlook its limitations. This amounts to offering faculty an elegant, well-tailored straightjacket suitable for specific pedagogical occasions.

At institutions where it is desirable to develop software for a wide range of courses and foreign languages, specialized CALL authoring tools have heightened value because of their transferability among instructors and departments. Although no tool will be fully useful for every foreign language because of graphics-related issues such as character display, with some modification most good interactions can be transferred from one foreign language to another. This will lead to formation of users groups that support collaboration among faculty and offer a channel for suggestions about how a given tool can be improved in succeeding versions.

Strategies for Selecting, Developing, and Adapting Software

Development of an institutional strategy for providing CALL courseware requires that decisions be made about all of the issues described above. The range of options can be viewed in three overlapping categories: implementation projects that involve the use of material developed by others; development projects that require substantial faculty effort to adapt existing software or to use CALL authoring systems and templates; and research projects that develop a whole application from scratch. A list of considerations for determining the right courseware strategy follows:

a. Use the work of others as a launching pad for any project; it provides cost effective shortcuts and helps the faculty build a critical perspective for future projects.
b. Decide how to adjust the tradeoff between offering leadership by creating new, leading-edge applications and benefiting from the cost-effectiveness of existing materials.
c. Consider a mixed strategy that includes both top-down major projects that serve the whole institution and bottom-up, faculty-initiated experiments that permit maximum creativity to bubble up; investment in a series

of small-scale projects that involve a wide range of faculty may yield more significant results in a shorter time than attempting to agree on and conduct a single massive and fully integrated project.

d. Without heavy involvement of the faculty most projects will fail.

e. For sophisticated projects that meet today's interface standards, a team approach involving technical specialists and faculty is essential; good software is developed by one or two dedicated teachers with a vision, assisted by a solid support team.

f. Software developers do what the current technology lets them do, as constrained by finite resources and finite programming expertise.

g. Good software matures like wine. It is not always clear at the start of a development process what the ideal design or even the ideal goals are. It is not always clear what the technology has to offer the pedagogy, because we have so few examples.

h. Good courseware has a generalizable form or it is probably not worth the costs.

i. Intra-institutional collaboration helps provide the faculty with a cadre of like-minded, dedicated developers even though only one or two experimenters may be present in a given department.

j. Inter-institutional collaboration seldom works for software development, although it can provide a basis for sharing materials and a springboard for new projects.

k. The benefits of creating CALL authoring tools militate against the creation of comprehensive tutorials and complex simulations, both of which may lack any generalizable form if they exploit the technology fully.

The Primary Issues in Software/Hardware Compatibility

Incompatibility results from the best efforts of software developers to use the full capabilities of the available development platform. CALL courseware exacerbates this tendency because of the need to exploit graphics, sound, and video to create complex learning environments. Even when multimedia issues are set aside, the gulf separating various hardware platforms has often meant that software developed on one platform either could not be adapted to another, or was extremely expensive to convert. This set of roadblocks prevailed throughout the 80s.

The 90s will be characterized by the evaporation of this problem in large measure. To the extent that major vendors support a graphical, window-like interface, event-driven interactions, and a pointing device, applications on their machines will or can resemble each other. The

concept of hypermedia has given rise to development tools that have many commonalities; in some cases conversion utilities allow partial translation of code from one machine to another. Sophisticated authoring systems may provide their own conversion utilities in order to make the system more attractive to a wide range of users. New generations of equipment from the same vendor must provide some level of backward compatibility with major software products of the past. Some platforms will provide software or hardware that can be used to run software created for an alien platform.

This encouraging general tendency will never account for all situations where hardware incompatibility is an issue. New hardware, especially peripheral devices, and new operating systems will catapult one platform into a new generation of software that cannot be matched on a competing platform for several years. Leading-edge applications will always suffer from a kind of quarantine; in some cases a specialized development platform will never be widely adopted and its software will come to be a museum piece as the machines it was written for become outdated. For these reasons, the temptation to develop CALL on highly experimental configurations of equipment should be viewed quite skeptically. Trailing-edge applications do have the virtue of hindsight.

Apart from machine compatibility, there is still the issue of interface compatibility. Institutions that use authoring tools developed elsewhere, or adapt/adopt courseware from other institutions, or encourage a wide variety of faculty projects will soon discover that students cannot move readily from applications to applications. The lack of national standards for something as simple as typing conventions for foreign characters or control schemes for video players cannot easily be remedied. These problems may sometimes outweigh the advantages of implementing a particular piece of courseware.

Here are a few last quasi-aphorisms, this time related to compatibility:

a. Use a platform in tune with national, rather than local trends.
b. Consider supporting a range of platforms, at least for experimental purposes, so that it is always possible to test out major advances in CALL courseware.
c. The ideal platform is a moving target; it is not always detrimental to remain with a successful platform several years beyond its heyday, since this conserves resources for the replacement generation of equipment.
d. Success in courseware development can make the curriculum dependent upon a given set of materials and machines; conversion to new equipment becomes more and more costly as the base of installed materials grows.

Conclusion

The many possible directions and opportunities for enhancing the foreign language curriculum via CALL cannot be reduced to a single path, approach, or project. In the 90s, administrators who want their institutions to remain in the forefront of instructional technology cannot afford to rely on their own intuitions and preferences. The watchword here must be diversity. Changing theories, creative faculty members, and productive courseware development strategies offer a threefold challenge to every institution:

First, instructional technology must allow for some degree of experimentation with theories and methods, no matter what the institutional policies may be. Even if methods are consistent across faculty and departments, they will change over time faster than new software can be developed. This dictates the development of software that can be adapted to changing needs.

Second, diversity is the key to capitalizing on the most creative energies of the faculty. Rather than crafting a policy that resembles a monolithic five-year plan for technology, visionary administrators must put into place incentive programs that lure faculty to solve their own problems using technology.

Third, a multifaceted development strategy will capitalize on the best available software and development tools, will discover new ways that technology can foster learning and acquisition in both the productive and receptive skills, and will keep pace with national developments on a variety of platforms and peripheral devices.

Author

JAMES P. PUSACK, Associate Professor, Department of German, University of Iowa. 103 Schaeffer Hall, Iowa City, IA 52242. Specializations: use of computer and video in foreign language instruction, language teaching methodology.

Leading the field in sociology and the related social sciences:

sociological abstracts (sa)

and

Social Planning / Policy & Development Abstracts (SOPODA)

Our subject specialists track the broad spectrum of theoretical and applied sociology from the more than 1,800 discipline-specific and related journals published in North America, Europe, Asia, Africa, Australia, and South America.

sa and **SOPODA** each offer you in-depth abstracts and precise indexing of timely journal articles and books, enhanced dissertation listings, and a bibliography of book reviews from the journals screened.

sa and **SOPODA** are available together on the **sociofile** CD-ROM and are hosted online by BRS, DATA-STAR, DIALOG, and DIMDI. Hardcopy subscriptions can be ordered from the address below.

The **sa** and **SOPODA** information products are supported by:

- Database-specific user manuals
- The latest journal coverage list
- The **sociofile** Quick Reference Guide and User's Handbook
- The **Thesaurus of Sociological Indexing Terms**
- Your Guide to Searching **sa** using a Personal Computer
- A professional workshop program

The **sa** family of databases — your fast track to the information you need, in the format you want.

sociological abstracts, inc.

p.o. box 22206 • san diego, ca 92192-0206
phone (619) 695-8803 / FAX (619) 695-0416 / Help Desk (800) 752-3945

Applied Language Learning
1991, Vol. 2, No. 2, pp. 77-91

Hardware for Language Training: Coping with Confusion

Michael D. Bush
United States Air Force Academy

Introduction

Research and development efforts are showing that there are advantages of using new interactive technologies in the language learning process. Effective use depends on the availability of powerful hardware and software. The choice of these two elements must of course rest on the functionality dictated by an adequate theory base for how people learn languages. The purpose of this article is to discuss what types of hardware are necessary to meet the requirements of such a theory base.

One of the more important questions that needs to be addressed is the determination of those technologies that show the greatest promise for the future. To approach this issue one must know what emerging hardware platforms are prominent today and will continue to be prominent in the future. Addressing platforms that will be useful in the years to come has a lot to do with forecasting. As was once said, "Forecasting is really a tough job, especially when it involves the future."

Paradigm Shifts

One conclusion drawn from our experience at the Air Force Academy is that to consider what new technology can do for the language learning process, one must seriously consider the need for a paradigm shift. Stated another way, efforts with new technologies should not simply automate the past. Although there is a normal tendency to do with new technology what was done with the old, effort must be expended on discovering what can be done with the new that was not possible in the past.

The concept of paradigms has been applied in many diverse fields. Thomas Kuhn (1970) outlined the notion in his *The Structure of Scientific Revolutions* and talked about how, contrary to popular belief,

science has not made slow and steady progress. Scientists have followed a time line of discovery that has consisted of plateaus and spurts. During the time that no revolutionary changes in thinking are made, commonly held theories dictate the types of conclusions that scientists can draw from observations they make in the world around them. Any apparently conflicting conclusions are interpreted in terms of commonly held beliefs until someone comes along who can break away from conventional wisdom and derive new conclusions.

A commonly cited illustration of the concept of paradigm shift has to do with the Copernican notion of an Earth that revolves about the sun. Most people of that time believed that the Earth was the center of the Universe. All they had to do was look at how the sun rose in the morning and set at night. To think otherwise was heresy!

Somewhat like Copernicus, there are a few popular consultants and speakers today who go out into the business world and say, "We really have to change the way we look at the world around us." Doing this, they say, is crucial because we always look at our new way of doing things in terms of the old. But, to make revolutionary rather than evolutionary progress requires that one take a step back and look at things differently, with a new perspective.

New Paradigms for Language Learning

The parallel to be drawn is that the best hardware platform for language learning most likely should not be specified in terms of the old way of teaching languages. It is necessary to think about the crucial questions: Should new technology be used to do things the same old way? Should the problems be examined in a different light and a tack taken that is different than the way things have been done to date?

Unfortunately, change is difficult for most people. For example, the great inventor and scientist Nikola Tesla came to this country specifically with the idea of going to work for Thomas Edison. Some of Tesla's initial findings while working in Europe had to do with the advantages of converting to the alternating current (AC) model for transmitting electricity. Edison was totally opposed to this particular technology and said it would never go anywhere. Of course, Edison was installing direct current (DC) generators to serve the homes of the Vanderbilts, the Astors, and the other "robber barons" of the day. Because the generators had to be close to the point where the electricity was to be used, he was installing direct current generators sometimes in the basements of the mansions, underscoring just how totally unwieldy his technology was. Tesla was convinced that this was not the way to distribute electrical power and, in fact, totally revamped the way that electrical transmission is carried out (Cheney, 1981).

The language teaching profession needs to be just as radical in examining how languages are taught and must establish which of the current methodologies that have been developed over the last several hundred years need to be changed. There are probably things that will remain the same, but there are probably others that will be different. The particular technology and the particular hardware available as tools for the teaching process probably constitute the greatest impetus for change. Of course, it is not necessary that there will be change. Change does not come about because new technology is available; rather change takes place because new technology makes possible new ways to do things that make sense with the emerging theory base.

This technology will invariably change the way we do business. The older technologies of writing and movable type were at one time revolutionary. After their invention, people did not do things the same way as they had before. Therefore, in implementing technology, it is essential to be attuned to the idea of paradigm shifts. One must accept the requirement to assume a totally different mind-set regarding the language teaching problem.

Hardware for Implementation

The first step in considering technology or hardware for language teaching is to consider an old computer industry axiom, a dictum for implementing hardware in other areas. As the saying goes, one should go out and find the software that one needs and then buy the computer that will run it. Although that might be very appropriate for areas where software exists, language learning does not really have a corpus of existing material from which teachers can select that which is most appropriate for their objectives.

This is an exaggeration to establish a point. Nevertheless, it is possible to consider what sort of functionality language learning hardware should have. Due to the sorts of things involved in the language learning process, this notion of functionality for hardware delivery platforms decides in advance the characteristics that must be implemented. Although the federal government spends tremendous amounts of money, there are budgets that must be respected. It does not make sense to purchase systems with capabilities that are not essential nor to omit putting in capabilities that are.

Three guiding principles will help make hardware implementation decisions:

- Software
- Functionality
- Price

Each of these areas is discussed below:

Software

Although perhaps a bit simplistic, the first thing a technology-based learning environment can accomplish is the presentation of information to the student. This also means that students can be placed into an environment where they can receive comprehensible input, something one particularly prominent theory says is vital for the language learning process. On the other hand, it is possible to ask questions of the learner. The level of interaction required by these questions is dependent of course on the capabilities of the software. Given sufficiently powerful hardware and adequate programming, it is also possible to have some form of intelligent interaction with the learner.

This description of software is as straightforward as possible and condenses a lot of concepts into two high-level notions. The instructional designer can cause the system to:

- Present information to the learner
- Ask questions of the learner

These two categories evoke the same idea of "expository" and "inquisitory" in Merrill's notation of "transaction shells" (Li & Merrill, 1989). The decision to select either or both of these two global types of alternatives will dictate the type of hardware that will be needed to effect a particular type of interaction with the learner. Just to present simple materials is fairly straightforward. After information has been presented to the learner, it is then important for the student to be able to respond and for the system to take action on that response. This interaction implies two additional software capabilities that the hardware must be able to support:

- Student Input
- Assessment and Presentation Control

Simply stated, the student should be able to respond, to input something into the system, and the system should be able to act on that information. If the system is to do more than respond to student input of simple predictable answers (fill-in-the-blank, multiple choice, or true-false questions) its software must have some level of knowledge about the language system being conveyed to the learner. Providing this capability requires artificial intelligence techniques that are only possible using more powerful hardware and running more sophisticated software than is necessary for simple presentation of materials.

This does not mean that full natural language processing is required. The state of the art is not there yet, especially for affordable delivery systems for language learning. There are, however, "smart"

things the computer can do and these should be explored (Schank, 1990).

As part of the assessment function, the system should be able to perform a couple of functions. First, based on the learner's input, the system should be able to assess at what level the student is, in the learning process. Second, based on the response of the student, the software should be able to direct activity to some other portion of the presentation. Most commonly held definitions of interaction will contain these elements.

Functionality

Presenting information to the learner requires consideration of the type of data that constitutes the information. Based on how languages are learned in general, it is possible to subdivide the process into dealing with the presentation of three types of data:

- Sound
- Images
- Text

Language itself, together with the way language is learned and used, are both related to these three elements that are each very much a reflection of the world in which language exists. Because these elements make up a substantial portion of the way people communicate, any system that is to teach all aspects of communication must, by definition, be able to handle these three types of data. Stated in other terms, the functionality that must be in instructional delivery hardware should reflect the world in which the subject matter resides.

Because sound is the essence of language, any language learning system must have the ability to deliver sound. Anything less is a poor reflection of reality.

Images are important, because spoken language is often a comprehensible process only when participants can see each other, examining each other's faces to see if appropriate messages are being communicated. When there is difficulty, phrases can be restated or clarification can be requested.

Text, as a reflection of the spoken word, has become very important in communication. Text was also the first type of data that early computers could deliver and remains important in the era of multimedia computers.

Falling Prices

With this fairly simple yet comprehensive list of required hardware functionality, it is necessary to consider the means that will

make the hardware widely available. Interactive videodisc has captured scattered interest throughout the industrial and military training markets; however, it has not had the success that many predicted at the outset. Stephen Jobs, the founder of Apple Computer, said to a group of educators two or three years ago that videodisc would never be successful because it lacked the resolution really needed to solve interesting problems. Jobs' assertion may or may not be true, but other factors have hindered videodisc's success. Specifically, hardware and materials development costs have been prohibitive.

There are two current developments, however, that will have a tremendous impact on the price and the type of hardware that will be available for instructional delivery systems:

- Advent of digital media technologies
- Impact of the mass market

The world is experiencing tremendous development in the area of miniaturization and increased power of digital, silicon-based technologies (Gilder, 1989). As an illustration of the changes these technologies are bringing, consider a discussion that took place between a group of interactive videodisc people and Bill Gates, the young chairman of Microsoft. At the first Microsoft Compact Disc-Read Only Memory (CD-ROM) Conference in 1986, this group of people, several of whom had been in videodisc dating back to the mid-70s kept telling Gates, "With digital compact disc you don't have the bandwidth that is possible with analog videodisc." Essentially, this statement means that compact disc (CD) does not have the ability to put as much information (visual data) on the screen as does videodisc. To paraphrase his response, Gates said not to worry about it. Everything that can be done with videodisc will ultimately be do-able in some digital format; moreover, once the video is in digital format, many exciting things will become possible within the computer. Gates, the "Crown Prince of computing technology," and the videodisc people were on different wavelengths. Gates was also saying that computers would be sufficiently powerful and memory readily available for the mass market to supply affordable delivery platforms on a wide scale.

To illustrate how this will happen, consider that the mass marketplace has embraced CD-Audio in a way that Sony and Philips never anticipated: people are buying CDs at such an incredible rate that local department and record stores have more CDs than they do 33-1/3 albums. These same manufacturers of CD-Audio are creating Compact Disc Interactive (CD-I), a new CD-based delivery system significantly smaller than videodisc. Because CD-I will benefit from the market penetration and manufacturing economies of scale of CD-Audio, this technology promises ultimately to be every bit as powerful and significantly less expensive than interactive videodisc. Sony has shown

prototype CD-I units that can be carried around to provide high quality audiovisual presentations in the palm of the hand.

By benefitting from this mass market success, computers can be implemented that are no longer specialized systems. It is no longer necessary to specify and acquire Military Specification (MIL-SPEC) devices to deliver language training, as was perhaps the case a few years ago. Mass market electronics in large part has been made possible through the miniaturization requirements of the space program and defense acquisitions, and it is now getting to the point where the Department of Defense can turn around and benefit from mass marketplace technology not even existing five or ten years ago.

The incredible reductions in price are having an astounding impact in the mass market. As an example, the computing power in household appliances surpasses that of mainframe computers from just a few years ago. At least for the next ten years, one must assume that the electronics industry will continue its advances and continue to provide increasingly powerful hardware at decreasing cost and in extraordinarily smaller packages. Thus, the mass marketplace is now having a significant impact on the way language learning systems can deliver the text, images, and sound necessary to support the learning process.

How Much Technology is Enough?

The concept of increased power for decreased cost raises an interesting question: How much technology is enough? In language learning terms the question becomes: How much technology is sufficient to meet the problems encountered in the language learning process? The increased power of the resulting hardware and software systems can be applied in providing improved and additional capabilities in three main areas:

- Fidelity
- Feedback
- Open Architecture

Although the low resolution of videodisc mentioned above might be a problem for some applications, there is a level of fidelity that is possible with videodisc that is not possible with other types of computer presentations. People are accustomed to seeing high quality images, video with high production value. They have a lot of money to spend on what they see in the movie theater and what they see on the television screen, a relatively high fidelity technology. They are used to high quality and will demand it.

During the initial work on interactive media at the Air Force Academy in the early 80s, computer systems were only capable of

representing the world as squared-off presentations and stick figures. Many researchers were convinced that this resolution was not going to be suitable for instruction. So, to answer the question of how much technology is enough, a system should be able to provide the same level of representational fidelity that people are generally used to in other aspects of their lives.

The delivery system must be powerful enough to give good feedback; in other words to present information to the learners, and then to have them be able to see an alternative presentation based on their input or something else based upon their decision or a decision made by diagnostic software. Platforms available for two or three years are sufficiently powerful for presenting information. Only now are platforms both affordable and sufficiently powerful to adequately raise the level of feedback that can be provided.

Open architecture is a hot topic in the microcomputer industry. This has application in technology-assisted language learning. To be useful, a computer must not only have adequate power to execute powerful programs quickly but must also be able to display several forms of data: text, graphics, audio, and video. To reduce instructional materials development costs, the formats of these data should be transportable across the platforms that are available in the marketplace. Transportability relates to standards, a very critical issue if technology is to realize its potential in learning.

To illustrate, when one goes into the grocery store and walks down the cereal aisle to make a purchase, it is not necessary to worry whether any particular cereal is going to be compatible with the milk at the other end of the grocery store, nor what type of place setting is available at home. The delivery system for the cereal is the bowl and spoon that will be used. The bowl can be porcelain or Corelle Livingware, and the spoon can be stainless steel or silver. All will get the job done.

Such simplicity is far from the rule with technology. Machines will soon transcend the issue of binary compatibility. It will no longer be necessary to take a program that runs on an MS-DOS machine and have that exact same program run on a MacIntosh or in a UNIX environment. Each of these environments has its strengths. The Department of Defense is currently heavily into MS-DOS and will be more and more into UNIX through the POSIX operating system. Considering the power that is becoming available in each of these environments, it will be a simple matter to require that the data (i.e., the graphics, audio, video, and text discussed above) used in a presentation on one machine be easily transportable to the other hardware environments. A slight conversion process might be necessary, but with proper software this will be an easy step for most of these computers.

How much technology is enough? It is absolutely essential that delivery platforms be powerful enough to make use of the standard data structures for each of the various essential data types. As an example, one can take a Microsoft Word document on an MS-DOS machine, save it in Microsoft's Rich Text Format (RTF), and then save the document onto a NeXT computer using Sun Microsystems' de facto standard Network File System (NFS). Accessing the document using NeXT's word processing software makes the computer think it is reading an Apple Macintosh file. About 98 to 99 percent of the formatting information is present after the conversion. Being able to do this with text documents is exactly what that open architecture should bring to the multimedia data structures and file formats that are useful for machine-based language learning materials.

Quantity of Computers

How many computers should be implemented in the learning process? Several years ago Jean-Jacques Servan-Schreiber in his book, *The World Challenge (Le Défi Mondial)*, talked about silicon as the primary component of dirt. Because silicon makes up the major component of microcomputers, there is no reason why they should not be dirt cheap! Observing what is going on in the world, such an observation is very indicative of the price/performance ratio that exists in the marketplace today. Jerry Pournelle, a science fiction writer and columnist for InfoWorld and Byte magazine, has formulated what he calls Pournelle's Law. He says that as the price of computers drops, essentially approaching zero, the ratio of computers to people will be at least one to one. This means that every person will have immediate access to at least one computer.

This availability has interesting implications for the planning and budgeting processes. In such an environment, what is the proper ratio of computers per person? Will one computer for three people or one for every person be the right number? At the Air Force Academy, as well as at the other service academies, and at many other universities, management has made the determination that students will have their own computers. In the early 1980s, I was on some of the initial committees at the Air Force Academy that looked at this issue, and I remember how we analyzed various options. Our charter was to look at how many computers would be appropriate in the mid-1980s. One per student seemed to be the answer, but this ratio seemed a bit ludicrous to many at the time.

But, there are a lot of reasons to implement technology at this ratio on a larger scale throughout the educational system. Computers are going to be cheap enough. The interesting things that can be done with them are such that users will want to turn to the computer to

accomplish myriad information-related tasks. Computers will become essential tools, not unlike the allegory evoked some years ago about the idea of a time-sharing pencil—When pencils were invented, schools could not afford them in great quantities. Those that were purchased had to be passed around.

Configurations

Within the Department of the Army, many organizations are working with what are called the Electronic Information Delivery System, "EIDS-class" machines. Seven or eight years ago, the Army carried out a needs analysis and specified the type of technology that would be useful to deliver information electronically. Putting video and digital data capability together on videodisc, the Electronic Information Delivery System (EIDS) came into being.

There are problems that arise with implementing technology based on specifications that have to go through a process controlled by the federal acquisition regulations. Following the completion of a needs analysis, specifications are created and published to obtain contractors' bids. Following a contract award, those not receiving the contract protest. When the complaints are finally settled and the contract is awarded, the supplier builds and delivers machines that have become outdated during the process.

The other services, nevertheless, have implemented machines that fall into the same category of capability. However, there are much more powerful configurations possible today that can deliver some interesting interactive videodisc instruction with this class of machines. Furthermore, much can be done with these computers to keep them useful well into the future.

There are questions that need to be raised and issues that need to be addressed immediately:

- How should an organization upgrade existing technology?
- How is it possible to maintain a certain level of compatibility during technology advances?
- How can existing videodisc installations benefit from the compact disc systems that are coming from the mass marketplace?

Looking beyond the EIDS-class machine, the Language Learning Center (LLC) at the Air Force Academy is moving very quickly toward acquiring the capability to do multitasking. Multitasking means that the computer can seemingly do several things at once. For example, a typical desktop machine today used for typing a document is probably

idle about 95 percent of the time in terms of overall processing capability. Consequently, an operator could probably feed data to the computer at a rate a thousand or perhaps even a million times faster than can be typed. Only a small percentage of the computer's capability is being used during such operations.

How can this excess capacity be put to use? In learning centers it is desirable to diagnose students' status, to find out where they are and where they should go next, all based on their individual differences such as interests and learning styles. It is also desirable to implement the most intelligent ways possible to respond to the student input as discussed earlier. This implies the use of fairly sophisticated technology in the learning center.

There are other types of activities, however, that are possible for students to accomplish on their own as computer-assisted study. The consumer electronics market is developing Compact Disc Interactive (CD-I), a technology that will make this possible. First announced in 1986, CD-I companies are saying that it is finally to be available during 1991. Industrial CD-I players have been available for a couple of years and have been implemented in such applications as information kiosks.

At the Japan Electronics Show in Tokyo, Sony recently demonstrated a CD-I prototype that is hardly larger than the CD that is inserted in the player. The unit has a flip-up screen, 4-inch color liquid crystal display, and controls for interaction. It fits in the palm of the hand, obviously allowing the user to take it anywhere. Although this type of machine will not have the power to handle sophisticated answer processing that is possible on machines in learning centers, it will present high quality video with interaction as well as give the student access to tremendous amounts of data.

The quantity of data a CD can hold is enormous. For example, a standard compact disc will hold the equivalent of 650 megabytes, an amount that would require over 1,500 floppy disks. In terms of paper, this comes to about 270,000 typed pages of information. If audio is placed on the CD, it is possible to have random access to 16 hours of high quality audio and only use a small portion of the total potential bandwidth of the CD. On a disc that contains the maximum quantity of audio, it is possible to place several hundred megabytes of data.

This audio can be combined with good quality graphics, and very soon, full screen, full motion video (at better than VHS quality). Manufacturers anticipate that these devices will follow the same remarkable price drops that CD-Audio has experienced during its life cycle. Consumer players are expected to start around $1000, dropping to a few hundred dollars within just a few years.

No doubt the sort of capability seen with the robots of the movie Star Wars will some day be possible. It will eventually be possible to have tutors such as C-3PO that will have the ability to converse. Thus,

long-range plans should evolve and eventually include provisions for students to have individual tutors that carry on a focused conversation. Although technology will not be at that point in the foreseeable future, there are significant changes coming that demand an intelligent migration path.

Migration Path

All along this growth path there will be a lot of exciting language learning activities that various technologies will make possible. This paper began with a discussion of forecasting. One aspect of forecasting is that there is always the tendency to overestimate the long term. Work at the Air Force Academy eight years ago was hopeful, but the things we anticipated then were nothing compared to what we see happening in the next two years. The changes that are now anticipated require solid planning functions to ensure that implementation is as smooth as possible and that technologies are implemented at the proper point in their life cycle and at the right juncture on the migration path.

There is no question that videodisc is today's multimedia technology. It is a powerful technology with exciting possibilities that has been implemented on a large scale in the language learning centers of each of the Service Academies. These centers are probably the three largest centers of their kind in use anywhere. Hundreds of students a day are using interactive videodisc to facilitate the language learning process. And they are enjoying it! This is not a technology that should be implemented sometime in the future. It is here today. It is true that there are problems with producing materials, but there are various hardware and software developments on the horizon that are going to help that process.

As a significant part of any migration path, developers and users of interactive videodisc need to be thinking about what it will mean to implement digital systems and data. In a move never seen with interactive videodisc, the microcomputer and consumer electronics industries are promoting CD-based delivery systems under the banner of "multimedia". The videodisc industry has never had the likes of Bill Gates, the $3 billion kid (chairman of Microsoft), or James Canavino, the President of IBM Entry Systems Division, or David House, the President of the Microcomputer Components Division of Intel, or John Scully, CEO at Apple, or Rod Canyon, CEO at Compaq Computers, proclaim that the future of any existing major industry is dependent on videodisc technology. They have made this statement about multimedia.

At the same time that the consumer electronics industry is supporting CD-I, several companies in the microcomputer industry (including Intel, IBM, and Microsoft) are supporting Digital Video

Interactive (DVI). DVI has full-screen, full motion video right now and the CD-I companies have shown what its video will look like, although the decompression chips to be included in the players are not yet available. Both technologies will most likely benefit from international standards that are being determined by the Motion Picture Experts Group (MPEG).

Because it is derived from CD-I and a part of the audio plans for DVI, CD-ROM Extended Architecture (CD-ROM XA) has great potential as a transition technology. It can deliver large quantities of digital data, 16 hours of audio, and lots of visual data; but it does not have full screen, full motion video. For those applications where videodisc is already present, it is a simple matter to add CD-ROM XA drives to the upgrade plans of system developers. The Air Force Academy will be installing 35 CD-ROM XA units into half of its IVD workstations.

Given the class of 80286-based machines into which they fall, the EIDS delivery systems have some capability to move into the future. By definition, EIDS is a Level IV system and consequently can handle digital data on the videodisc and thus, theoretically, does not need the additional capability provided by XA. The difficulty, however, is that these digital programming and data files are non-standard. Based on the concept of open architecture discussed above, producing and using data and program files in the EIDS format locks developers out of the mass marketplace, denying them access to all of the tools that will be produced to support CD-I, DVI and CD-ROM XA.

Summary

To summarize, educators will have access to powerful, interactive technology in the future. The exact form it will take is not known, but we can safely say that it will be digital. We can also safely conclude that some form of compact disc will be present in the delivery system. Whether the consumer electronics people are successful in putting a computer into the CD player or the computer people put a CD player into the computer, either way we are going to have multimedia capability. To arrive at this conclusion only requires that one look at *Newsweek*, *Time*, *Business Week*, and numerous other publications to see the new buzz word, "multimedia." Multimedia will give us, through the mass marketplace, capabilities that we could only dream about a few years ago. In fact, the computer companies are taking steps that show they feel their survival might well be dependent on how well they implement multimedia.

Why is this? The computer people are convinced that multimedia is going to help them sell computers, and this is a great motivation. There are 80 million or so microcomputers out there right now, but this

represents only a 20 percent penetration of the total potential market. The captains of the microcomputer industry want to sell more. They want one on every desk and in every home. However, to do that they must make the machines easier to use, a task they feel multimedia will help them accomplish.

The consumer electronics industry wants to continue on the wave of its success with CD-Audio. They feel like they've reached a plateau with this technology. But by putting visual information, full-screen, full-motion video, in interactive packages they will be able to get into consumers' pocketbooks as well. That is the bottom line.

Whether it comes through the consumer electronics or microcomputer companies, one way or the other language educators are going to have the capability to do some very exciting things. Making the software happen requires some new ways to store and recall visual and auditory information using digital systems.

Another event in the very near future will be stylus input or the ability to actually recognize handwriting. The devices to be produced will be so small there will not be room for a keyboard to type on. Manufacturers will produce these devices with sensitive screens for the stylus to make selections, somewhat like a mouse.

Now all of this is wonderful, but one cannot ignore the fact that all humans have a basic resistance to change. People are basically conservative. Marshall McLuhan said that, "Innovation to the holders of conventional wisdom is not novelty, but annihilation."

Some language teachers think along these lines and are afraid that their jobs are threatened. This is not so. As hardware and software become more powerful, the job of teachers will change, but their services will be required into the foreseeable future. They will do things differently in the classroom, but they will still be teaching and students will still be learning. Technology will be a new tool to be brought to bear on the teaching/learning process. Experience at the Air Force Academy has shown that one of the biggest challenges is teaching teachers how to do things in new ways. Hardware advances will continue, as will advances in software. A far greater challenge will be the political implications of helping teachers learn how to do their job differently, a challenge far greater than the technical ones.

References

Cheney, M. (1981). *Tesla: A man out of time*. Englewood Cliffs, NJ: Prentice-Hall.

Gilder, G. (1989). *New microcosm: The quantum revolution in economics and technology*. New York: Simon and Schuster.

Kuhn, T. S. (1970). *The structure of scientific revolutions (2nd ed.)*. Chicago: The University of Chicago Press.

Li, Z. & Merrill, M.D. (1989). Transaction shells: A new approach to courseware authoring. *Journal of Research on Computing in Education*.

Schank, R. (1990). Teaching architectures. *Research Report of The Institute for the Learning Sciences*. Northwestern University, Evanston, Illinois.

Servan-Schreiber, J. J. (1980). *Le défi mondial*. Paris: Fayard.

Author

MICHAEL D. BUSH, Chief, Learning Technology Division, Department of Foreign Languages, *United States Air Force Academy*, CO 80810. Specializations: technical and pedagogical implications of using technology, innovative ways to produce materials.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editors encourage the submission of manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on 8½ x 11 inch paper, double-spaced, with ample margins. Where feasible, manuscripts are preferred on disc (DOS or XEROX Viewpoint). Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and if necessary notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

Review Process

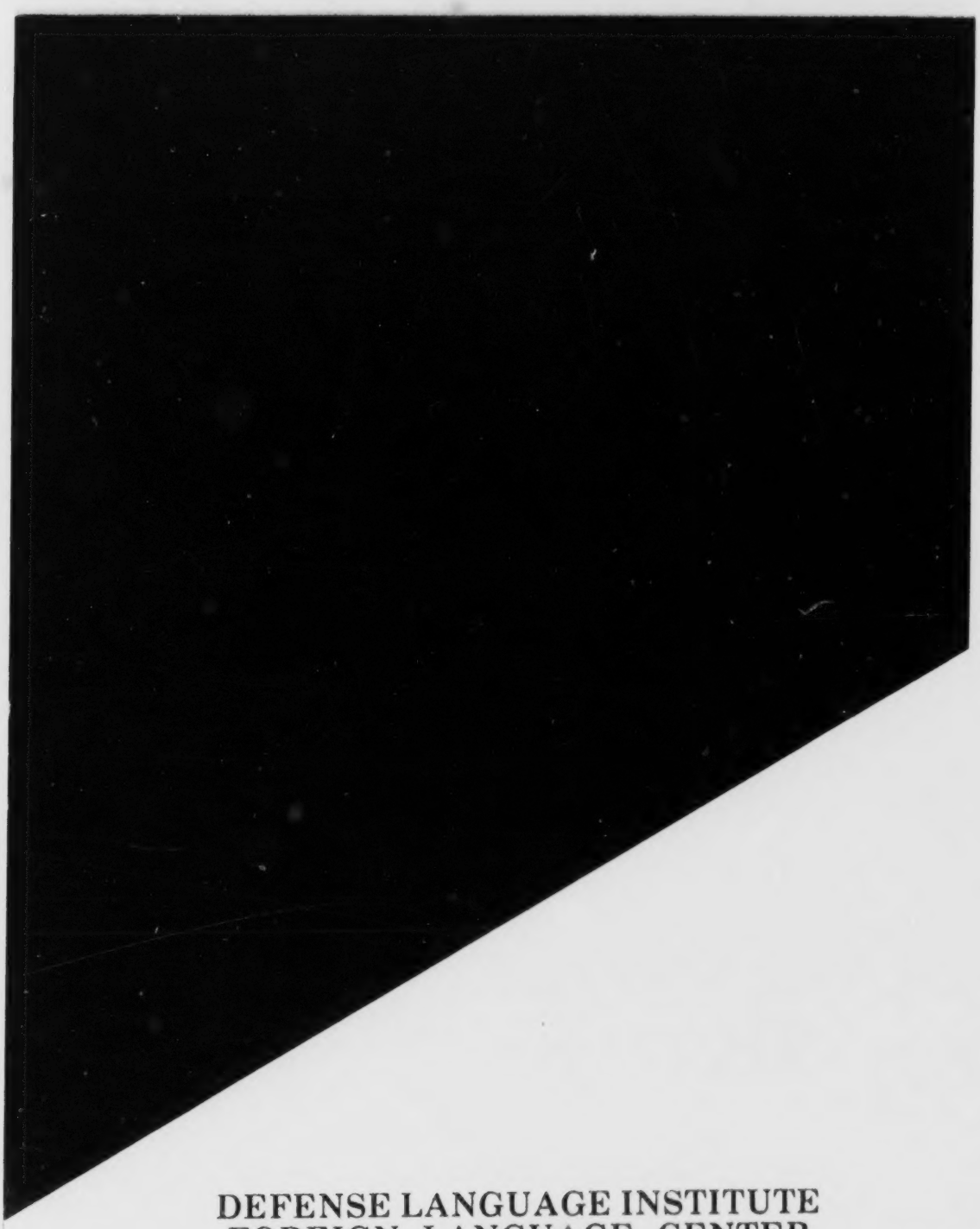
Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

GPO 685-001/42441

PIN 055711 — 000



DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

306

Applied Language Learning

VOLUME 3 • NUMBERS 1 & 2
1992

307

Applied Language Learning

Volume 3 • Numbers 1 & 2

1992

Defense Language Institute
Foreign Language Center

308

APPLIED LANGUAGE LEARNING

PB 65 - 92 - 1 & 2
US ISSN 1041-679

The mission of *Applied Language Learning (ALL)* is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Professional Bulletin 65-92, *Applied Language Learning* published semi-annually by the Defense Language Institute, Foreign Language Center, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

GORDON R. SULLIVAN
General, United States Army
Chief of Staff

Official:

MILTON H. HAMILTON
Administrative Assistant to the Secretary of the Army
03108

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to DLIFLC.

Applied Language Learning is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Book rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster — Send change-of-address information to: *Applied Language Learning*, Office of the Provost, Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. Telephone: (408) 647-5638 or DDN: 878-5638.

APPLIED LANGUAGE LEARNING

PB 65 - 92 - 1 & 2

1992 • Volume 3 • Numbers 1 & 2

Lidia Woytak
Editor

Editorial Board

Kathleen M. Bailey	Head, TESOL/ESL Program, Monterey Institute of International Studies
David P. Benseler	Editor, *The Modern Language Journal*
John L. D. Clark	Dean, Program Evaluation, Research, and Testing, DLI
Ray T. Clifford	Provost, DLI
Martha Herzog	Dean, Curriculum and Instruction, DLI
Betty Lou Leaver	Dean, Central European School, DLI
John A. Lett, Jr.	Director, Research, DLI
June K. Phillips	Director, Tennessee Foreign Language Institute
Gary D. Walter	Librarian, DLI

DLI Staff

Editorial Coordinator, Erika Malz
Production Editor, Jack Boyd
Cover Design and Graphics, Barney Inada
Layout Consultant, Annemarie Frazier

APPLIED LANGUAGE LEARNING

PB 65 - 92 -1 & 2

1992 • Volume 3 • Nos. 1 & 2

CONTENTS

1 Language Learning Strategies: Crucial Issues of Concept and Definition
Rebecca L. Oxford and Andrew D. Cohen

37 Factors Relating to Learning a Foreign Language among High- and Low-Risk High School Students and Students with Learning Disabilities
Leonore Ganschow, Richard Sparks, James Javorsky, and Jon Patton

65 The Utility of Objective (Computer) Measures of the Fluency of Speakers of English as a Second Language
Deborah Rekart and Patricia A. Dunkel

87 From the Editor: A Note of Appreciation

88 Information for Contributors

Applied Language Learning
1992, Vol. 3, Nos. 1 & 2, pp. 1-35

Language Learning Strategies: Crucial Issues of Concept and Classification

Rebecca L. Oxford
University of Alabama

Andrew D. Cohen
University of Minnesota

PURPOSE

The dual purpose of this article is:

1. to note seven key problems in language learning strategy concepts and classification systems as they relate to empirical research in the area of second and foreign language development; and
2. to suggest ways to resolve these problems and thereby increase unity, coherence, and meaningfulness of language learning strategy research.

This article is not intended as a comprehensive review of language learning strategy studies; for such reviews, see Oxford (1992/1993), Oxford and Crookall (1989), Skehan (1989), and O'Malley and Chamot (1990). The goal is instead to highlight the key conceptual and classificatory problems in the language learning strategy research area and to recommend solutions.

DEFINITION AND SCOPE

Language learning strategies are defined as steps or actions taken by learners to improve the development of their language skills. These strategies have the power to:

(a) increase attention essential for learning a language;
(b) enhance rehearsal that allows linkages to be strongly forged;
(c) improve the encoding and integration of language material, and
(d) increase retrieval of information when needed for use (Oxford, 1990b; Mayer, 1988).

Language learning strategies affect "the way in which the learner selects, acquires, organizes, or integrates new knowledge" (Weinstein & Mayer, 1986, p. 315).

However, language learning strategies do not just deal with the cognitive domain. Strategy use is also intended to "affect the learner's motivational or affective state" (Weinstein & Mayer, p. 315). In addition to the cognitive aspects of language learning, language learning strategies involve social (interpersonal, interactional), affective (emotional, motivational, attitudinal, and personality-related), and metacognitive (planning- and evaluation-related) aspects.

Language learning strategies are relevant at any point on the "learning-and-acquisition continuum" described by Brown (1987). That is, these strategies are used in formal classroom situations where language is presented systematically, as well as in less formal environments where language is picked up naturally without intentional instruction.

QUICK SUMMARY OF RESEARCH ON LANGUAGE LEARNING STRATEGIES

During the last two decades, language learning strategy research has burgeoned. Researchers have shown that successful learners tend to apply strategies in an orchestrated way relevant to their own needs and the characteristics of the task (see, e.g., Cohen, 1990; O'Malley & Chamot, 1990; Oxford, 1990a, 1990b, 1992/1993; Vann & Abraham, 1989; Wenden, 1991; Wenden & Rubin, 1987). According to some researchers, less successful language learners sometimes do not even know what strategies they use, or they are aware of just a few noncommunicative and mundane strategies, such as translation, rote memorization, and repetition (Nyikos, 1987). Yet sometimes less effective learners are indeed aware of using a wide range of strategies, but they employ these strategies in a random, almost desperate manner, not keyed to the task at hand (Vann & Abraham, 1989).

Strategy training studies in the foreign and second language fields have frequently been successful but not consistently so (Cohen, 1990; Chamot & Kupper, 1989; O' Malley & Chamot, 1990; Tang, 1990). Careful observation shows that methodological problems—too short a period for strategy training, too easy or too difficult a task, lack of integration of strategy training into regular class activities, lack of perceived relevance of strategy training, and inadequate pretraining assessment of learner's needs—might have obscured potentially important findings in some cases.

CONCEPTUAL AND CLASSIFICATORY PROBLEMS TO BE SOLVED

From the profusion of studies devoted to language learning strategies, recently summarized by Oxford (1992/1993), one might believe that this research area is fully coherent. However, this coherence is something of an illusion. Seven serious conceptual and classificatory problems exist in the language learning strategy research field. These problems do not cancel the significance of research findings in this realm, but they do underscore an urgent need to clarify assumptions about categorizations as soon as possible, so that future research can be more sound.

Problems in the language learning strategy research area include the following:

1. Distinctions between "strategy" and "tactics" are unclear or nonexistent.
2. Researchers often disagree about whether learning strategies are conscious or unconscious.
3. Many different criteria are used for classifying language learning strategies—the "moving target" syndrome.
4. Researchers argue about *learner* strategies vs. *learning* strategies—or what contributes to learning a language.
5. The concept of "strategic competence" in the well-known communicative competence framework is too narrow.
6. Many studies fail to show the conceptual link between language learning strategies on the one hand and learning styles, other personality-related variables, and demographic factors on the other.
7. So far, it is impossible to discern appropriate language learning strategy classifications for natural settings or technology-assisted instruction.

Each of these problems is discussed below, along with potential solutions.

Problem 1: Distinctions between Strategy and Tactics Are Unclear or Nonexistent

The word *strategy* comes from the ancient Greek word *strategia*, meaning generalship or the art of war. Today the military meaning of strategy involves the optimal management of troops, ships, or aircraft in a planned campaign. Strategy is "the conduct of conflict . . . the set of policies used for the conduct of conflict . . . the policy tool for planning aggression against opponents . . . the organized deployment of resources to achieve specific objectives against competition from rival organizations" (James, 1984, p. 12).

Another important word is *tactics*. This plural noun emerged from the Greek verb *tasso* (to arrange). It refers to the art "of disposing military or naval or air forces, especially in actual contact with (the) enemy; . . . (a) procedure calculated to gain some end, by skillful devices" (Fowler & Fowler, 1977). Thus, in military parlance strategy is a broader concept than tactics, although both concepts involve a goal- or problem-orientation, a sense of conflict, a set of actions, and some degree of conscious intentionality.

Von Clausewitz, the military strategist, made clear the distinction between tactics and strategy: "Tactics is the art of using troops in battle: strategy is the art of using battles to win wars" (James, 1984, p. 15). To translate Von Clausewitz's distinction into language learning terms, we might say that

> *learning tactics* represent the short-term art of using *specific behaviors or devices* (such as simplifying, inferencing, categorizing, asking for verification, creating personal examples, using positive self-talk, relaxing with music, setting goals, deciding to avoid distractions, arranging for conversation partners, and staying in a conversation through miming) to support one or more major learning strategies during day-to-day learning situations. *Learning strategy* is the long-range art of learning more easily and effectively by using *major clusters of behaviors* for forming concepts and hypotheses, testing hypotheses, personalizing linkages, embedding material in long-term memory, understanding one's affective state, managing the learning process, and producing language while lacking adequate linguistic knowledge.

Notice that in language learning, the strategy-and-tactics duo has shed its conflictual, warlike features but has retained the other

characteristics: active, conscious movement toward a goal. See Dunkin and Biddle (1974) for a discussion of strategy and tactics in education.

Schmeck's Strategy-Tactic Distinction

The distinction between strategy and tactics has been maintained by educational psychologists such as Schmeck (1988). Schmeck considers strategy to be "a general approach or plan of the learner," as well as "a higher-level cluster of learning tactics that work together to produce a uniform learning outcome." His examples of strategies include *conceptualizing, personalizing, and memorizing*. To Schmeck, a tactic is "a specific activity of the learner," an observable operationalization of a broader strategy. For instance the *conceptualizing strategy* can include tactics such as categorizing, comparing, contrasting, hierarchically organizing, abstracting, and networking ideas. The *personalizing strategy* can encompass tactics like self-referencing, generating examples, translating into personal language and images, and linking new information with prior personal experience. The *memorizing strategy* is operationalized through tactics such as repetitive rehearsing of information, using mnemonics, and encoding verbatim. (Schmeck feels that the memorizing strategy is a weak one when compared with the conceptualizing and personalizing strategies and that the memorizing strategy, unlike the other two, is associated with neurotic or insecure personality features. We believe that the memory strategy is not a neurotic manifestation but a sometimes sensible, if rather limited, behavior used in early stages of language learning for coping with large amounts of vocabulary to be learned.)

Seliger, Faerch and Kasper, and Marton on Strategies and Tactics for Language Learning

Seliger (1982) poses the distinction between strategies and tactics for language learning. To Seliger, strategies are sets of biologically determined, constant, abstract cognitive functions used to acquire knowledge; three kinds of strategies for language acquisition exist, according to Seliger: hypothesis testing, simplification, and generalization. In contrast, tactics are lower-level activities carried out in response to the immediate, local context of language acquisition. Similarly, Faerch and Kasper (1980) define strategies as plans for controlling the order in which a sequence of operations is to be performed. Process, according to Faerch and Kasper, consists of the operations (tactics) themselves.

Marton (1983) agrees in principle with Seliger's distinction between strategies and tactics, but reduces Seliger's three types of

strategies to one, hypotheses testing, and subsumes the other two—simplification and generalization—under it. Marton adds that "where tactics activate underlying basic strategies they will lead to successful long-term acquisition," but when the learner's filter or the learning environment "prevents a given tactic from triggering basic strategies, only language-like behavior or short-term acquisition can take place" (p. 314). Thus, strategies and tactics must support each other for the most effective long-term language development.

Later Seliger (in Larsen-Freeman & Long, 1991) adds to the strategy-tactic split by positing a new distinction between macro-tactics and micro-tactics. Macro-tactics, according to Seliger, result in situations where learners can obtain data (and may be similar to Rubin's 1981 "indirect strategies"), whereas micro-strategies provide direct input for learning (and may be related to Rubin's category of "direct strategies").

Where Strategies and Tactics Become Confused

Mayer (1988) unintentionally provides an excellent example of the confusion between strategies and tactics. He lists the following as learning strategies: underlining key ideas, outlining, and putting ideas into ones own words. These behaviors would seem to be tactics that operationalize broader strategies. The first two might reflect a strategy like conceptualizing or finding the central meaning, and the last might reflect a personalizing strategy.

Oxford's (1990b) language learning strategy list contains a number of strategies that, in retrospect, might much more accurately be termed tactics, such as analyzing expressions, looking for conversation partners, and using relaxation techniques to improve concentration and retention.

Too Many Fuzzy "Synonyms"

Larsen-Freeman and Long's (1991) landmark book on second language acquisition research reflects the confusion by appearing to equate learning strategies with all of the following: learning behaviors, cognitive processes, and tactics (p. 199). Wenden (1987, 1991) notes that these terms are all used synonymously: learning strategies, techniques, potentially conscious plans, consciously employed operations, learning skills, cognitive abilities, processing skills, problem-solving procedures, and basic skills. Oxford (1990b) adds that learning strategies are also equated with thinking skills, thinking frames, reasoning skills, tactics, and learning-to-learn skills.

What's Wrong with Equating Strategies and Tactics?

> Most of our interventions in learning strategies training programs are actually at the specific level of tactics, and we often fail to give sufficient attention to whether or not the tactics that we teach fit together to produce a coherent learning outcome (Schmeck, 1988, p. 173).

Thus, strategy training often focuses on small behaviors that may or may not be congruent or effective for improving learning.

Another problem in equating strategies and tactics is that doing so produces strategy lists that are overly long, atheoretical, and filled with mixed levels of behavior (some broad, like elaboration and inferencing, and some narrow, like taking notes and summarizing). "Hypertaxonomizing" then becomes a peril, as reflected in some of the work of Oxford (1990b) and others.

Solution to the Problem

Identifying broad strategies and distinguishing them from tactics would simplify our understanding of language learning strategies significantly. Language learning strategy researchers are advised to develop and agree upon a smaller number of broad strategies with their related tactics listed underneath each strategy. In this way, greater conceptual clarity would be obtained and O'Malley, Chamot, Stewner-Manzanares, Russo, and Kupper's (1985a) complaint about confusion in hierarchic relationships among strategies would be significantly reduced in intensity. *Table 1* offers a possible outline of strategies and tactics for language learning. This table will be referred to throughout the article.

TABLE 1
A Possible Organization of Language Learning Strategies and Their Supporting Tactics Useful for All Four Language Skills

---

1. *Strategy = forming concepts and hypotheses (building declarative knowledge).*
   Tactics = simplifying through transfer, simplifying through overgeneralization; using inferencing to figure out the meaning in the absence of adequate vocabulary and grammar; categorizing, comparing, contrasting, hierarchically organizing, analyzing, reasoning deductively, abstracting, creating networks, propositions, and schemata (conceptual frameworks). (See Faerch &

Kasper, 1980, 1983; Ellis, 1986; Oxford, 1990b; Schmeck, 1988.) This group contains tactics that have often been known until now as *cognitive strategies* and *compensatory strategies for listening and reading*.

2. *Strategy = testing hypotheses.*
   Tactics = testing hypotheses by means of intake analysis, assessing correctness of feedback, consulting experts, asking for verification or correction. (See Faerch & Kasper, 1980, 1983; Ellis, 1986; Seliger, 1980.) This group contains *tactics* that have often been classified until now as *metacognitive strategies* and *social strategies*.

3. *Strategy = personalizing mental linkages.*
   Tactics = referring to oneself, creating personal examples, using personal language and images, linking new information with personal experience and prior knowledge—also called elaboration. (See Schmeck, 1988; Oxford, 1990b.) This group contains *tactics* that have often been contained until now in lists of *cognitive strategies*.

4. *Strategy = embedding new material in long-term memory.*
   Tactics = repetitive rehearsing of information; using memory devices like imagery, keyword technique, rhyming, acronyms, physical response; imitating; encoding by rote or verbatim; formal practicing; and functional practicing. (See Schmeck, 1988; Oxford, 1990b.) This group contains tactics that have often been termed until now *memory strategies* or *cognitive strategies*.

5. *Strategy = understanding one's affective state.*
   Tactics = using positive self-talk, discovering ones personality type and learning style, assessing ones feelings, using a diary, relaxing by means of music or breathing, using humor. (See Brown, 1987; Dansereau, 1988; McCombs, 1988; Horwitz & Young, 1991; Oxford, 1990b.) This group contains tactics that have often been termed until now *affective strategies*.

6. *Strategy = managing ones learning process.*
   Tactics = setting goals, evaluating ones progress toward goals, checking comprehension, testing ones knowledge, planning for language tasks, deciding in advance to pay attention to specific aspects of input, deciding in advance to pay attention to the task and avoid distractions, arranging for practice partners, looking for practice opportunities, previewing, looking for the organizing principle, organizing the learning environment and materials. (See

O'Malley & Chamot, 1990; Oxford, 1990b). This group contains *tactics* that have often been called until now *metacognitive strategies*.

7. *Strategy = producing oral or written language while lacking adequate linguistic knowledge.*
   Tactics = *communicative achievement tactics* such as gesturing, miming, paraphrasing, literally translating, substituting, restructuring, approximating the message, coining words, switching to the native tongue, waiting for the item to come to mind; and *communicative reduction tactics* such as avoiding use of certain rules, speech acts, and functions. (See Tarone, 1980, 1981; Ellis, 1986; Oxford, 1990b.) Most of these strategies can be used for both writing and speaking. This group contains *tactics* that have often been called until now *compensatory strategies* or *communication strategies*.

Assumption #1

Strategy use involves some degree of conscious awareness on the part of the learner. When a behavior is used *unconsciously* (e. g., production behaviors described by Ellis, 1986, for automatizing the learner's existing linguistic knowledge through initial simplification, later complexification and self-checking), it is *not* a strategy because conscious awareness is not present, according to our definition of strategy.

Assumption #2

All of the strategies and tactics above contribute to learning, either directly or indirectly. For example, Tarone (1980) and Littlewood (1979) show that Strategy 7 enhances learning indirectly. Therefore, it is artificial to bracket off "learner strategies" (such as communication strategies) from "learning strategies," as some researchers have done.

Assumption #3

Although strategies are sometimes categorized as *procedural knowledge*, i.e., knowledge of procedures (O'Malley & Chamot, 1990, p. 231, after Anderson, 1980; Pressley et al., 1990), nevertheless, strategies are frequently used to create *declarative knowledge*, i.e., knowledge of facts, concepts, and relationships. Learning strategies can aid learners in making initial links between new material and stored information, strengthening existing links, and retrieving information through declarative networks when needed (Mayer, 1988).

Problem 2: Researchers Disagree about Whether Learning Strategies Are Conscious or Unconscious

In the learning strategy area, there are clearly distinct points of view about conscious awareness as an inherent feature of strategies. One group of researchers states that consciousness is an essential characteristic of strategies, and another group disagrees.

Consciousness as an Essential Feature of Strategies

The ancient Greek definition and most modern uses of the term "strategy" imply consciousness and intentionality. Weinstein and Mayer (1986) state that learning strategies must be intentional and hence conscious on the part of the learner, because learning strategies have a goal: facilitation of learning. Likewise, Rabinowitz and Chi (1987) suggest that strategies must be conscious; if they are performed automatically, they are no longer strategies. Chamot (1987b) asserts, "Learning strategies are intentional on the part of the learner" (although O'Malley and Chamot's 1990 book weakens that assertion, as shown below). Bialystok (1990) notes that all definitions of strategies involve problematicity, consciousness, and intentionality, although a contradiction appears (see below).

The conscious awareness involved in strategy use is highlighted by Miller, Galanter, and Pribram (1960) in their well-known book *Plans and the Structure of Behavior*. These researchers show that plans—equivalent to learning strategies—are largely based on a match or mismatch between the stimulus (incoming information) and the learner's stored information structure (known as a schema, or a configuration of interrelated features that define a concept). When incoming information does not agree with stored information, the learner responds by initiating plans of action to restructure the old information (similar to Piaget's [1952] operation of "accommodation"). When incoming information does agree with stored information, the learner does not have to restructure the old information, but instead simply expands the existing schemata to encompass the new information (similar to Piaget's operation of "assimilation"). The learner's plans (strategies) are intended to achieve conceptual coherence and self-control. These plans specify the sequence and organization of behavior patterns, the behaviors that are appropriate under particular conditions, the standards which must be achieved, and the consequences of reaching or falling short of these goals (Mischel & Mischel, 1977). Thus, according to Miller, Galanter, and Pribram, behavior shows meaningful action toward goals; strategies or plans are conscious units in the psychological structure of behavior.

In his book on language learning, Cohen (1990) defines strategies as learning behaviors which are consciously selected by the learner. He considers the element of *conscious choice* important because in his view it gives a strategy its special character. Thus, strategies are moves which the learner is at least partially aware of, even if full attention is not being given to them. There must be some level of awareness that a strategy is being used. If a learner's behavior is totally unconscious, then it would simply be referred to as a process, not a strategy. For example, a learner may use the behavior of skimming a portion of text in order to avoid a lengthy illustration. If the learner is at all conscious of why the skip is taking place, then it would be a strategy. Otherwise, it would just be seen as an unconscious process.

Consciousness as a Non-Essential Feature of Strategies

On the other side of the fence stand those who assert that a strategy need not involve consciousness or awareness on the part of the learner. As we have seen above, Bialystok (1990) notes that all definitions of strategies involve consciousness and intentionality; however, she seems to contradict these definitions by suggesting that strategies *cannot* have consciousness as a criterion: "There is no means of distinguishing plans that lead to strategic speech from those that do not by virtue of consciousness" (p. 5).

Similarly, O' Malley and Chamot (1990), who strongly emphasize Weinstein and Mayer's statements about learning strategies being intential and thus conscious (p. 43), nevertheless state that use of language learning strategies is part of procedural knowledge, which—though it might start out as conscious—is characterized chiefly by automaticity and lack of conscious awareness (pp. 52, 231). "Procedural knowledge includes mental activities such as problem solving, language reception and production, and *using learning strategies* . . . Production systems [are] the processes by which procedural knowledge is stored in long-term memory . . . and, once learned, will tend to be *executed rapidly and without awareness*" (p. 231, italics added). Thus, O'Malley and Chamot underscore the strong possibility that language learning strategies can be unconsciously performed.

One researcher suggests that some language learning strategies involve conscious/behavioral processes and other language learning strategies involve unconscious/psycholinguistic processes (Tarone, 1980). The borderline between conscious and unconscious is not always clear, however.

Some researchers use the term "potentially conscious" to refer to language learning strategies (Faerch & Kasper, 1980; Wenden & Rubin, 1987). This term implies that language learning strategies might be, in some instances, unconscious as well. Yet there is little clarity about

what makes a strategy conscious in some instances and unconscious in others for that same learner or different learners.

There is even less clarity about whether strategy use moves from consciousness to unconsciousness (or vice versa). One could assert that the aim should be to heighten the conscious awareness of strategy use by language learners; on the other hand, one could also assert that the aim should be to make language learning strategy use so natural and practiced that it becomes fully automatic and unconscious.

A Related Issue: Strategy Consciousness versus Language Input Consciousness

The concern for consciousness about the use of language learning strategies is sometimes confused with more general consciousness about second or foreign language learning. There is currently a concern for having language learners become more aware of the *language material* they might wish to learn. Ellis (1990), for example, sees consciousness-raising as a legitimate goal of form-focused instruction—i.e., as a means to the attainment of grammatical competence. Schmidt (1990) asserts that "noticing" is the necessary and sufficient condition for converting input to intake. He goes on to suggest that "paying attention" is probably facilitative in language learning and may be necessary for adult learners if they wish to acquire grammatical features. Scovel (1991) emphasizes attention as "central to the entire process of second language acquisition" and as "the neuropsychological mechanism that promotes or prohibits acquisition" (pp. 3, 5). *These kinds of conscious attention refer to paying attention to the language material and are different from being aware of using language learning strategies*—although both consciousness of the language material and consciousness of strategy use are beneficial. The issue of "potential consciousness" that is found in discussions of students' awareness of learning strategies does not appear full force in discussions of students' awareness of language material.

Solution to the Problem

It is important to take a stand on whether language learning strategies are conscious or unconscious. We assert that language learning strategies are indeed conscious behaviors undertaken to improve language learning. If strategies are unconsciously and automatically used, then explicit strategy training makes little or no sense. Research shows that strategy training, when conducted well, can improve students' strategy use by making learners more fully conscious of their current behaviors and by bringing to consciousness additional

strategy options for learners to practice and adopt if useful (Cohen, 1990; Oxford & Crookall, 1989).

Empirical research on language learning strategies (Green, 1991) provides some justification for making consciousness an essential criterion for these strategies. Green reports that his advanced language learners often have significantly lower strategy use than intermediate language learners, and that intermediates use strategies significantly more than do beginners. Thus, strategy use in Green's study might appear to be curvilinear, with intermediates using language learning strategies far more than advanced and beginning language learners. One might speculate that advanced learners might have automatized their learning behaviors, so they might not use or need language learning strategies as much as do intermediates; and beginners might not have yet developed a large, conscious, and frequently tapped repertoire of strategies. For clarity's sake, it seems useful to call conscious learning behaviors strategies but to avoid using that term for automatized, unconscious learning behaviors.

Teachers and students must keep in mind differences between (a) the learner's consciousness of strategy use and (b) the learner's awareness of language material to be learned. When one talks about strategy awareness, one is referring to the learner's understanding of his or her own strategy applications—how he or she takes in new language material, encodes it, and transforms it to make it usable for actual communication. When one discusses general language consciousness, one is referring to the learner's awareness of language input. Yet there might be a link between these two concepts. Paying attention to language input may involve the conscious use of language learning strategies, such as strategies for attending to input, strategies for determining the purpose of input, and strategies for linking input with existing knowledge. Therefore, raising awareness about the language material might also enhance awareness about the use of language learning strategies.

Problem 3: Many Different Criteria Are Used for Classifying Language Learning Strategies— the "Moving Target" Syndrome

Many contrasting criteria are employed for categorizing language learning strategies by the same or different investigators. This creates a "moving target" syndrome in which a given behavior appears to dart from one category to another, depending on the inclinations or epistemological understanding of the researchers. Some of the main criteria for categorizing language learning strategies are shown below. Note that a great deal of overlap exists in the sets of criteria. The main thrust(s) of a given strategy classification system are highlighted.

Finally, many of the behaviors that are viewed below as *strategies* might more aptly be called *tactics*, as explained earlier in this paper.

1. Strategies *used by successful language learners*

   This is a criterion that has been adopted by Naiman, Frohlich, and Todesco (1975); Rubin (1975, 1981, 1987); and Stern (1975, 1983). (The opposite criterion, strategies used by unsuccessful language learners, has been employed as a classifier by Vann & Abraham, 1989). A variety of language learning strategies (and some tactics), as well as a range of personality variables that are not really behaviors, is included in these systems.

2. *Cognitive (intellectual) and metacognitive (planning-organizing-evaluating) functions*

   The many related but not identical language learning strategy taxonomies of O'Malley, Chamot, and their research group (e.g., Chamot & Kupper, 1989; Chamot & O'Malley, 1987; O'Malley & Chamot, 1990; O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985a; O'Malley, Chamot, Stewner-Manzanares, Russo, & Kupper, 1985b) have concentrated on cognitive and metacognitive functions. These taxonomies have only very lightly touched the surface of social strategies and especially affective strategies, the latter of which "are of less interest in an analysis such as ours which attempts to portray strategies in a cognitive theory" (O'Malley & Chamot, 1990, p. 44). Some broad strategies, such as elaboration, are mixed in with what we view as tactics, such as outlining.

3. *Declarative and procedural knowledge*

   The cognitive information-processing, declarative-procedural metaphor of cognition by Anderson (1980; see also Sternberg, 1990, for a comparison with other current "metaphors of mind") was adopted by O'Malley and Chamot (1990) and by Faerch and Kasper (1983). Declarative knowledge relates to facts, ideas, propositions, networks, concepts, whereas procedural knowledge relates to procedures like using learning strategies. This computer-based metaphor leaves out almost all affective aspects of language learning. The strategy systems of O'Malley and Chamot as well as Faerch and Kasper are quite different, despite their use of the declarative-procedural contrast. No distinction is made between strategy and tactic by O'Malley and

Chamot, although Faerch and Kasper distinguish between strategy and operation (which seems like tactic).

4. *Directness or indirectness*

   Rubin (1981, 1987) identifies language learning strategies that purportedly contribute directly to language learning (direct strategies, such as cognitive and metacognitive strategies) and those that contribute indirectly to language learning (indirect strategies, such as social strategies and communication strategies). The means for discerning the directness or indirectness were not fully clear. Oxford (1990b) uses the direct-indirect split in a different way, with directness or indirectness relating to whether the target language itself is directly or indirectly involved in the use of a given strategy. The strategy-tactic split is not involved in these frameworks.

5. *Whole-person characteristics*

   On the assumption that the learner is a multifaceted human being, not just a cognitive-metacognitive information-processor, Oxford (1985, 1990a, 1990b) has created a series of strategy systems that focus on a wide range of features, with the most recent system including large numbers of strategies (some more rightly viewed as tactics) in each of six categories: memory-related, cognitive-processing, compensatory, metacognitive, affective, and social.

6. *Strategies versus tactics for cognition*

   Seliger (1983) and Marton (1983) distinguish between broad strategies (e.g., hypothesis-testing) and smaller tactics with a focus on cognitive behavior in language learning. They avoid any discussion of affective, social, and metacognitive aspects. They present some of the few language learning systems that clearly divide learning behaviors into strategies and tactics.

7. *Explicitness and type of knowledge*

   Bialystok (1978) and Ellis (1990) have developed strategy frameworks based on the degree of explicitness (or implicitness) of knowledge and the kind of knowledge (e.g., linguistic or world knowledge; form-focused and meaning-focused knowledge). Language learning strategies are then related to these factors.

8. *Communication strategies*

Tarone (1977, 1981, 1983), Faerch and Kasper (1983), and Ellis (1986) have paid attention to strategies (often really tactics) that compensate for inadequate linguistic knowledge in a *speaking* situation. Tarone and Yule (1989) have expanded this concept to include compensatory behaviors in a *listening* situation. Though these are not often called "learning strategies," Tarone (1981, 1983) makes it clear that learning can and does take place when they are used. (Note: Production/reception strategies are employed for using the learner's existing linguistic knowledge automatically, according to Faerch and Kasper (1983), and are not seen as learning strategies; see Ellis, 1986.) Sometimes these communication strategies are called *compensatory strategies* (see Ellis, 1986; Kellerman, 1991; Littlewood, 1984; Oxford, 1990b; Poulisse, Bongaerts, & Kellerman, 1984; Poulisse, 1989b; Tarone & Yule, 1989). Poulisse (1989) shows that these strategies can be applied either holistically or analytically for listening and speaking (although writing and reading are not addressed); Oxford (1990b) demonstrates that these can be applied to all four skills.

9. *Language skill or learning mode*

Cohen (1990, 1991) focuses on how language learning strategies (or in some instances, what we might term tactics for language learning) fit into different language skills or learning modes, such as oral production, writing, reading comprehension, vocabulary learning, and attention enhancement.

When so many different criteria are used for categorizing language learning strategies, it is natural that the same strategy is classified differently by the same or different investigators. This makes reading the strategy research a very confusing activity and makes it almost impossible to summarize results across studies. The following three examples suggest some of the complications that occur when the same strategy (or actually, tactic) is classified different ways:

1. *Asking questions for clarification or verification* is classified as both a social strategy (e.g., Chamot, 1987a; Chamot & O'Malley, 1987; Chamot & Kupper, 1989; O'Malley et al., 1985b) and as a cognitive strategy (O'Malley et al., 1985a; Rubin, 1987). A parallel confusion exists for the behavior known as cooperation.

Probably asking questions for clarification and verification does have both cognitive and social aspects.

2. *Monitoring one's own errors* is usually categorized as a metacognitive strategy (see, e.g., Oxford, 1990b; Chamot & Kupper, 1989; O'Malley et al., 1985a). However, it is listed as a cognitive strategy by Rubin (1987), who claims that this strategy contains both cognitive and metacognitive elements. Stern (1983) classifies monitoring as an academic (explicit) learning strategy.

3. *Seeking practice opportunities* is called a metacognitive strategy (Oxford, 1990b) and a social strategy (Rubin, 1987). The probable assumption underlying Rubin's classification is that practice takes place with other people, while Oxford assumes that the learner might practice alone and/or with others. Also, classification of this behavior might depend on whether the emphasis is on the *seeking* or on the *actual conduct* of the practice opportunities.

The system in *Table 1*, which focuses on broad strategies with related tactics shown to support the strategies, avoids many of the double and triple classifications found across different researchers. It does so by focusing on the very basic clusters (strategies) of intential behaviors as the organizing principle within a whole-person framework, with related tactics flowing directly from the strategies.

Strategy categorization conflicts do not just occur *across* investigators; they also arise for a single researcher or single group of researchers. An individual or a group often publishes somewhat different lists of language learning strategies at different times. This typically results from the evolution of a series of research studies, an evolution which is not usually explained very well (or at all) to research consumers such as teachers, students, or administrators.

The lack of concordance in strategy categories and concepts, while confusing even for the most diligent and intelligent reader, is actually related to differences in research methodology, and research methodology is a reflection of researchers' goals. Some of the identifiable goals include: analyzing the causes of linguistic error, describing a good language learner, defining learning processes, and validating teaching methods. Clearly, the goals of the research on cognitive and metacognitive strategies in the language area and the rarity of research on affective and social strategies may suggest that most language learning strategy researchers value the former more than the latter. The imbalance indicates where the investigators' chief

concerns lie but implies little about the actual importance of these strategies or their contribution to language learning.

It is not surprising that strategy research methods vary greatly, and these methods strongly affect the categories and numbers of language learning strategies found in any study. For instance, if investigators use only a classroom observation scale, they are likely to miss all the unobservable concept- and hypothesis-formation behaviors that learners use. If researchers employ student surveys which omit learning strategies for understanding emotions and motivation, they will discover nothing about learners' use of such strategies. If scientists employ individualized think-aloud procedures without any student-to-student or student-to-teacher interaction, a host of interpersonal, interactive strategies may fail to appear.

Solution to the Problem

Although it would be a slight exaggeration to say that the field of language learning strategies has "one strategy classification system for each strategy researcher," this description is not too far from the mark. Investigators should come up with one or more acceptable sets of criteria for categorizing language learning strategies. *Table 1* is an intentional attempt to do this, based on the concept of learning strategies for the "whole person" and the distinction between strategies and tactics. With so many different criteria used for establishing strategy systems as foundations of major research studies, it is predictable that language learning strategy investigations would often be confusing and contradictory. This intolerable situation, if it endures, will continue to hinder the language learning strategy field from achieving its scientific and educational potential. Differences in strategy classifications between groups of researchers need to be discussed and, when possible, resolved publicly in print, so that the significance of the research can be better understood.

Moreover, researchers need to be more explicit about how and why their minds change regarding classifications of language learning strategies. Multiple language learning strategy typologies from the same researchers (including the authors of this article) need to be explained as they evolve. In addition, researchers should be very clear in their research design. They should indicate any limitations on their results related to the way the research design is established. Better yet, researchers should use multimethod, multifactor studies that involve many different data collection and analysis techniques, both quantitative and qualitative. In addition, a combination of longitudinal and cross-sectional research would give a better picture than either one by itself (O'Malley & Chamot, 1990).

Strategy studies would benefit from the cooperation of groups of researchers in different locations, especially in different countries (as is now happening with a major language learning strategy study led by the first author, and encompassing Japan, Puerto Rico, Egypt, and mainland US).

Problem 4: Researchers Argue about Learner Strategies versus Learning Strategies—or What Contributes to Learning a Language

A tricky semantic problem in the language learning strategy field is the distinction between learner strategies and learning strategies (Ellis, 1986). Wenden (1991) and Wenden and Rubin (1987) refer to *learner strategies* in the titles and texts of their books. O'Malley and Chamot (1990), Oxford (1990b), and Cohen (1990) prefer to use the term *learning strategies* in their books. Tarone (1980) and Ellis (1986) use both terms for distinct purposes.

Learner strategies, according to Tarone (1980), are simply any strategies used by language learners. Tarone identifies three kinds of learner strategies: *learning strategies*, *production strategies*, and *communication strategies*. To Tarone, learning strategies are the actions by which the learner processes the second language input to gain linguistic knowledge. The goal of learning strategies is to assist the learner in language acquisition or learning.

Purportedly separate from these learning strategies (and thus viewed as "non-learning" learner strategies) are *production strategies* and *communication strategies*. These two groups are not called learning strategies because they supposedly do not directly contribute to learning. The goal of production and communication strategies is language use rather than learning.

1. *Production strategies*, according to Ellis (1986), include strategies employed *unconsciously* to enhance the automatic use of language knowledge already acquired; examples are simplification at early stages of language learning, complexification at more advanced stages (Littlewood, 1979), planning speech before producing it, and correcting speech afterwards (Seliger, 1980). (Note: It is not clear how planning and correcting can take place without some degree of consciousness, however.)

2. *Communication strategies*, rather narrowly defined as strategies for overcoming roadblocks in spoken communication although not in written communication, fall into two groups, according to Ellis (1986), who based his research on Faerch and Kasper (1983): *reduction*

> *strategies* that involve giving up part of the communication goal through avoiding use of certain rules, speech acts, and functions; and *achievement strategies* that involve keeping the communication goal but compensating in some way, as in code switching, substituting, paraphrasing, word coining, restructuring, gesturing, and asking for help. Tarone (1983) notes three kinds of communication strategies: *paraphrasing* (approximation, word coinage, circumlocution); *borrowing* (literal translation, language switch, appeal for assistance, mime); and *avoidance* (topic avoidance, message abandonment).

Kellerman (1991) argues that most attempts at setting up taxonomies of communication strategies (and other compensatory strategies) are based on conflicting and inexplicit criteria and thus have doubtful psychological validity. Nevertheless, some of these taxonomies are widely used and reprinted (Scarcella & Oxford, 1992; Ellis, 1986).

Although production strategies and communication strategies are used by learners, and are therefore viewed as part of the range of learner strategies, as noted above, there are researchers who assert that these cannot be called learning strategies, because they do not have learning as a primary goal (e.g., O'Malley & Chamot; Selinker, 1972). We believe that this is a false dichotomy, because any strategy—such as a communication strategy—that allows the learner to keep participating in communication will be a potential enhancer of learning (Littlewood, 1979). Tarone (1981, 1983) states that communication strategies can indirectly enhance learning. Tarone argues that it is virtually impossible to distinguish between learning strategies and communication strategies for three reasons: (a) it is difficult to assess the real purpose for which the person uses the strategy; (b) the motivation might be both learning and communication simultaneously; and (c) even when the purpose is communication, learning often occurs anyway.

Solution to the Problem

There is no advantage for language learning strategy researchers to continue to quibble over whether a particular strategy directly or indirectly affects learning, as long as there is *some effect on learning*. There is no benefit to students or anyone else when investigators eliminate from strategy lists all purported indirectly-contributing strategies, such as the so-called communication strategies. In fact, if such "learning" versus "non-learning" distinctions are made, the confusion about what contributes to learning will only be perpetuated.

As shown in *Table 1*, it is feasible to include communication strategies in a broad organizational scheme of language learning strategies. Strategy 7 (producing oral and written language although lacking adequate linguistic knowledge) encompasses tactics that have traditionally been known as communication strategies or compensatory strategies; these contribute to learning by allowing students to continue to communicate even without all the requisite linguistic resources. (Production strategies are not shown in this taxonomy because Tarone (1980) and Ellis (1986) view them as unconscious and automatic.)

Problem 5: The Concept of "Strategic Competence" in the Well Known Communicative Competence Framework Is Too Narrow because It Is Usually Restricted to Speaking

The communicative competence framework of Canale (1983) and Canale and Swain (1980) has had a major effect on the foreign and second language field. This framework—which stresses grammatical, sociolinguistic, discourse, and strategic competences—is widely viewed as a four-part "prescription" for language instruction. Omaggio's (1986) book on proficiency-based instruction highlights the utility of the communicative competence framework, as does Scarcella and Oxford's (1992) Tapestry Approach to language learning and teaching.

However valuable the overall Canale and Swain communicative competence theory, the *strategic competence* aspect of this theory is limited, referring only to the learner's ability to use a small group of speaking-related strategies to compensate for inadequate grammar or vocabulary. These are the same strategies (really tactics) that Tarone (1980) and Ellis (1986) call *communication strategies*. For example, the learner uses gestures or circumlocution when the precise word is not known in a conversation. Thus, competence in using *compensatory speaking strategies* is the *only* type of strategic competence mentioned in the Canale and Swain framework of communicative competence.

More recently, Tarone and Yule (1989) have expanded strategic competence to encompass the use of communication strategies not just by the speaker but also by the listener. Thus, guessing and interpretation by the listener are at last admitted to the realm of strategic competence. While Tarone and Yule's expansion is a sign of progress, it is still not fully adequate because it suggests that compensatory strategies are not relevant to the major skill areas of reading and writing. Oxford (1990b) points out that the same general principles apply to compensating for missing knowledge in reading and writing, just as in speaking and listening. For instance, writers, like speakers, must often find ways to express themselves when they do not know all the relevant words; thus they use circumlocution and word coinage to get the point across. Readers, like listeners, frequently

encounter words they do not know and then have to infer the meaning from the context.

Ellis (1986) provides a good explanation of language learning strategies, but he often appears to restrict those strategies to the skill of speaking (and possibly the accompanying skill of listening). He provides examples of language learning strategies for formulaic speech and for creative speech:

1. *Strategies for formulaic speech* include pattern memorization, pattern imitation, and pattern analysis (see also Seliger, 1982; Fillmore, 1986; Krashen & Scarcella, 1978).

2. *Strategies for creative speech*, in which the learner produces entirely novel sentences, are based on Faerch and Kasper (1980, 1983) and include:
   a. *strategies for establishing interlanguage rules* (hypothesis formation, simplification, inferencing, and hypothesis testing); and
   b. *strategies for automatizing interlanguage knowledge* (formal practice and functional practice).

Clearly, these language learning strategies refer just to oral-aural skills but could easily be expanded to reading and writing.

While early reference to strategic competence as a component of communicative language use (Canale & Swain, 1980) put the emphasis on "compensatory" strategies, that is, strategies used to compensate or remediate for a lack in some language area, Bachman (1990) provides a broader theoretical model for viewing strategic competence. There is an assessment component whereby the person sets communicative goals, a planning component whereby the person retrieves the relevant items from language competence and plans their use, and an execution component whereby the person implements the plan. After finishing, the learner may again perform an assessment to evaluate the extent to which the communicative goal was achieved.

Within this broader framework, it may still be the case that a fair number of strategies are, in fact, compensatory. Learners often omit material because they do not know it when put on the spot, or produce different material from what they would like to, with the hope that it will be acceptable in the given context. They may use lexical avoidance, simplification, or approximation when the exact word escapes them under pressure, or possibly because they simply do not know the word that well or at all. Yet probably other strategies that fall under the rubric of strategic competence in Bachman's model are not compensatory.

As is the case with any theoretical model, learners may make differential use of the components of this model when performing

specific communicative tasks. For example, there are learners who do not assess the situation before starting to communicate, and because of this they might violate certain sociocultural conventions. Likewise, there are learners who plan out their utterances before producing them while others would just start talking immediately. Recent research involving the use of verbal report directly after the performance of oral role-play interaction is beginning to obtain data regarding the extent of strategic assessment and planning actually taking place before the execution of apologies, complaints, and requests (Cohen & Olshtain, in press).

Solution to the Problem

Many people who adhere to the well-known Canale and Swain communicative competence framework have not realized the need to expand the typical usage of the concept of strategic competence to encompass compensatory and non-compensatory behaviors in all four skills: listening, reading, speaking, and writing. Any strategy framework needs to be broad enough to include language learning strategies for reading and writing as well as for listening and speaking. The language learning strategy framework in *Table 1* addresses all four skills and deals with both compensatory and noncompensatory strategies that might be part of strategic competence. Tarone and Yule's, Oxford's, and Bachman's attempts to expand the strategic competence concept need to be disseminated widely and considered well.

Problem 6: Many Studies Fail to Make the Conceptual Link between Language Learning Strategies on the One Hand and Learning Styles, Other Personality-Related Variables, and Demographic Factors on the Other Hand

Learning strategies do not operate by themselves, arising *de novo* every time to meet the need of a new and specific learning situation. Instead, they are directly tied to the learner's underlying learning styles (general approaches to learning) and other personality-related variables (such as anxiety and self-concept) in the learner (Brown, 1991). They are also related to demographic factors such as sex, age, and ethnic differences (Oxford, 1989; Politzer, 1983).

Schmeck (1988) underscores the need to understand learning strategies in the context of *learning styles*, which he defines as the expression of personality specifically in the learning situation and which he divides into three categories: deep, elaborative, and shallow styles. Schmeck also exhorts researchers to view learning styles and learning strategies in the context of *general personality factors* such as the following: personality type (introvert/extrovert, reflective/impulsive,

field independent/dependent, which have often been viewed by other researchers as aspects of learning style, as shown in Oxford, 1990a), self-confidence, self-concept self-efficacy, creativity, anxiety, and motivation (intrinsic/extrinsic). His system is as follows:

1. *Deep learning style*, related to the *learning strategy of conceptualizing* (with *tactics* such as categorizing, comparing, hierarchically organizing). This style is connected with a stable, introverted, internally controlled, reflective, field independent personality that exhibits high self-confidence and intrinsic motivation. *Learning outcomes* are synthesis, analysis, evaluation, subsumption, schema development, and theory development. (Note: *Table 1* lists some of these outcomes, such as analysis, as tactics.)

2. *Elaborative learning style*, associated with the *learning strategy of personalizing* (with *tactics* such as self-referencing, generating personal examples, using personal images). This style is linked with a stable, extroverted, internally controlled, impulsive, field dependent personality that displays high self-confidence and a combination of intrinsic and extrinsic motivation. *Learning outcomes* are application, personal growth, personality development, social development, and understanding of people.

3. *Shallow learning style*, linked to the *learning strategy of memorizing* (with *tactics* such as repetitive rehearsing, using mnemonics, encoding verbatim). This style is associated with a neurotic, externally controlled, anxious personality that shows low self-confidence and extrinsic motivation combined with fear of failure. *Learning outcomes* are descriptions of what was studied and literal reproduction.

A learning strategy disembedded from personality-related factors is "only a short-term prop for learning," according to Schmeck (1988, p.179, italics added). The highest learning performance and the best satisfaction are associated with congruent, well integrated packages of learning strategies along with personality-related factors, including learning style. It is easier to modify learning tactics than to alter underlying factors such as learning strategies and personality-related variables (including learning styles). However, modifying learning tactics without dealing with underlying factors can create severe dissonance in the learner. For some learners, it might be necessary to modify learning style, especially the shallow learning style.

Cronbach and Snow (1977) in their classic work on aptitude-treatment interaction stated that learning styles may moderate either the effectiveness of certain learning strategies or the effectiveness of the procedures used to teach those strategies to students. Thus, an understanding of students' learning styles may be essential for using learning strategies effectively or for providing useful training in improvement of learning strategies. Likewise, Ellis (1990) shows that instructional input is filtered through any given students' learning style.

Similarly, Pressley and associates (1990) emphasize that contemporary models of thinking (e.g., Baron, 1985; Brown, Bransford, Ferrara & Campione, 1983; Sternberg, 1985) are multidimensional, including strategies (procedural knowledge), knowledge about those strategies and about one's own thinking processes (metacognition), knowledge about the world in general (the knowledge base), motivational beliefs, and overall learning style. These components operate interactively. Therefore, research that ignores any of these factors is likely to result in only a partial picture of learning and thinking (Costa, 1990).

> Whereas older models of thinking emphasized the "hardware" and/or a few "critical" cognitive processes (e.g., Atkinson & Shiffrin, 1968), these recent conceptualizations address the cognitive, metacognitive, and social-emotional aspects of thinking that both affect classroom functioning and can be altered using instruction. Thus, recent models of thinking are both more complete and more educationally relevant than their predecessors (Pressley et al., p. 8).

Naiman, Frohlich, and Todesco (1975) note that addressing the affective demands of language learning (related to emotions, attitudes, motivations, beliefs, and personality factors) is an important characteristic of "good language learners," along with more cognitive strategies. Bialystok (1978, 1981) shows that the learner's attitudes, world knowledge, and linguistic knowledge are important in strategy use. Politzer (1983) demonstrates how important it is to link strategy use with affective variables such as career motivation and with demographic factors like ethnicity and sex. Stern's list of strategies (1983) emphasizes social and affective strategies as well as active planning strategies and formal academic strategies. In a list of characteristics of successful language learners, Rubin (1987) notes psychological characteristics (e.g., risk-taking, tolerance for ambiguity, and willingness to appear foolish) and social learning strategies (e.g., seeking practice opportunities) along with communication strategies and cognitive strategies. Brown's books (1987, 1991) highlight the

crucial affective elements in language learning, and Horwitz and Young (1991) point out the role of language learning anxiety and cite language learning strategies to overcome such anxiety.

However, these wide-ranging variables have been largely ignored in most empirical language learning strategy studies in the last decade, which have often failed to gather, analyze, or report crucial personality-related, social, and demographic data about the subjects. Outside of the language field, researchers have clearly recognized the affective domain and its important role (see Weinstein, Goetz, & Alexander, 1988). Unfortunately, many language learning strategy studies in the 1980s have looked at cognitive and metacognitive strategy use but have frequently ignored factors such as motivation, beliefs, attitudes, anxiety, learning style, world knowledge, sex, and ethnicity—a situation that is both surprising and scientifically foolhardy.

The lack of concern about linkages between strategies and the affective (personality-related, motivational, and emotional) and social sides of the learner is reflected in many language learning strategy taxonomies, which usually include either no affective and social strategies (e.g., Bialystok, 1978, 1981; Carver, 1984; Marton, 1983; Seliger, 1982) or include only a small number of affective and social strategies dwarfed by large numbers of cognitive and metacognitive strategies (e.g., Chamot & O'Malley, 1987; Chamot & Kupper, 1989; O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985a and 1985b). Most language learning strategy survey and think-aloud procedures (described by Oxford, 1990b; Cohen, 1987) fail to include sufficient ways for students to report much use of affective and social strategies.

The situation is gradually improving through the "personological" work of some strategy investigators (see, e.g., Ely, 1989; Ehrman & Oxford 1988, 1989, 1990; Galloway & Labarca, 1991; Oxford 1990a, 1990b; Sutter, 1989).

Solution to the Problem

Language learning strategy research could benefit from a greater range of variables in the future. Researchers could routinely gather information on sex, ethnicity, age, degree of language learning experience, world knowledge, motivation, anxiety, beliefs, attitudes, and learning style—along with data on the learning environment and teacher variables. This would make investigations much richer and far more informative. Just as importantly, it would also allow researchers to develop an appropriate theory base for strategy use. More comprehensive selection of variables would no doubt necessitate larger studies, which might cause researchers to work together more

frequently rather than conducting smaller-scale, noncomparable studies. This would be an extremely helpful move.

In addition, language learning strategy taxonomies and assessment instruments (such as surveys, diaries, interviews, and think-aloud procedures) could be expanded to provide broader coverage of strategies that involve understanding of social and affective aspects of learning. Instead of giving very limited or nonexistent attention to strategies for managing these aspects, taxonomies and assessment tools could begin to reflect the true importance of these factors (Baron, 1985; Brown, 1987; Costa, 1990; Dansereau, 1988; Horwitz & Young, 1991; McCombs, 1988; Oxford, 1990b; Weinstein, Goetz & Alexander, 1988).

Any theory of how language learners learn cannot fail to include these aspects along with more overtly cognitive factors. Cognition is profoundly affected by social and affective variables such as emotions, attitudes, motivations, anxiety, social interaction, and personality type, as well as demographic variables such as age, sex, and ethnicity.

Problem 7: So Far It Is Impossible to Discern Appropriate Language Learning Strategies Classifications for Natural Settings or Technology-Assisted Instruction

Few studies have been done with language learning strategies in natural settings outside of the formal classroom. Underwood (1975) calls for the external validity of laboratory processes (i.e., learning strategies as studied in the classroom or in psychological laboratories) to be supported by evidence in more natural settings. Information on the use of strategies in a plethora of settings—inside and outside the ordinary classroom—is important for the development of a truly comprehensive system of strategies. Much more information is required on the use of strategies for gaining language skills in non-classroom environments (Oxford & Crookall, 1989).

Special learning situations, such as computer-assisted language learning (e.g., videodisc, intelligent tutoring systems, and computerized simulation) might also provide important data about the strategies people use. Technology provides a ready means of monitoring and recording strategies.

Solution to the Problem

Researchers should take up the challenge of gathering language learning strategy data in informal settings. Diary studies, strategy surveys, interviews, and think-aloud procedures are relevant tools for collecting such information outside of language classes. For example, learners could use the keyword technique for learning vocabulary in natural settings and then keep a diary about how effectively this

technique worked for them; all we have is laboratory data on this technique at the moment. Comparisons of strategy use between formal classroom settings and informal out-of-class settings would be extremely valuable for elucidating what really happens in language development. Because computers offer such detailed and helpful data on students' movements and behaviors, they are an ideal means for gathering strategy data in computerized language learning situations.

SUMMARY AND CONCLUSIONS

The authors of this paper have discussed seven problems in the research field of language learning strategies. Most of the problems deal directly with classification and conceptualization of these strategies. We have cited the problems in detail (as manifested in our own classification systems as well as the systems of others), and we have also offered potential solutions. Further discussion among researchers—and among users of the research—would certainly help identify additional solutions to these difficult conundrums. Rather than avoiding discussion of the problems (and thus exacerbating them), investigators should come to grips with these difficulties. Strategy researchers are encouraged to come together in a spirit of cooperation to provide and enact solutions to the problems mentioned in this article, recognizing that the language learner and the language teacher are our most important clients. In order to do this, researchers may have to set aside certain theoretical and methodological presuppositions that they have been making.

References

Anderson, J. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.

Atkinson, R. C. & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2). New York: Academic Press.

Baron, J. (1985). *Rationality and intelligence*. Cambridge: Cambridge University Press.

Bialystok, E. (1979). A theoretical model of second language learning. *Language Learning, 28*, 69-83.

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *Modern Language Journal, 65*, 24-35.

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Oxford: Basil Blackwell.

Brown, A. L., Bransford, J. P., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markham (Eds.), *Handbook of child psychology, Vol. 3: Cognitive development* (pp. 177-206). New York: John Wiley & Sons.

Brown, H. D. (1987). *Principles language learning and teaching* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Brown, H. D. (1991). *Breaking the language barrier: Creating your own pathway to success.* Yarmouth, ME: Intercultural Press.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt (Eds.), *Language and communication.* London: Longman.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1,* 1-47.

Carver, D. (1984). Plans, learner strategies, and self-direction in language learning. *System, 12*(2), 123-313.

Chamot, A. U. (1987a). The learning strategies of ESL students. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning.* Englewood Cliffs, NJ: Prentice-Hall.

Chamot, A. U. (1987b, March). The power of learning strategies. *Ohio Bilingual-Multicultural Update,* 6-11.

Chamot, A. U. & Kupper L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals, 22*(1), 13-24.

Chamot, A. U. & O'Malley, J. M. (1987). The cognitive academic language learning approach: A bridge to the mainstream. *TESOL Quarterly, 21*(2), 227-249.

Cohen, A. D. (1977). *Successful second-language speakers: A review of research literature.* (ERIC Document Reproducrion Service No. ED 163 753)

Cohen, A. D. (1983). Introspection about second language learning. *Studia Anglica Posnaniensia, 15,* 149-156.

Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 82-95). Clevedon, UK: Multilingual Matters.

Cohen, A. D. (1990). *Language learning: Insights for learners, teachers, and researchers.* New York: Newbury House/HarperCollins. (Now Boston: Heinle & Heinle).

Cohen, A. D. (1991). Strategies in second-language learning: Insights from research. In R. Phillipson, E. Kellerman, E. L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research* (pp. 107-119). Clevedon, Uk: Multilingual Matters.

Cohen, A. D. & Aphek, E. (1981). Easifying second language learning. *Studies in Second Language Acquisition, 3*(2), 221-313.

Cohen, A D. & Olshtain, E. (in press). The production of speech acts by ESL learners. *TESOL Quarterly.*

Cohen, A. D. & Robbins, M. (1976). Toward assessing interlanguage performance: The relationship between selected errors, learner's characteristics, and learners' explanations. *Language Learning, 26*(1), 45-66.

Costa, A. L. (Ed.) (1990). *Developing minds: A resource book for teaching thinking.* Alexandria, VA: Association for Supervision and Curriculum Development.

Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Halstead.

Dansereau, D. (1988). Cooperative learning strategies. In C. Weinstein, E. Goetz, & P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 103-120). New York: Academic Press.

Dunkin, M. J. & Biddle, B. J. (1974). *The study of teaching.* New York: Holt, Rinehart, & Winston.

Ellis, R. (1986). *Understanding second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (1990). *Instructed second language acquisition.* Oxford: Basil Blackwell.

Ely, C. (1989). Tolerance of ambiguity and use of second language learning strategies. *Foreign Language Annals, 22*(5), 437-445.

Faerch, C. & Kasper, G. 1980. Processes in foreign language learning and communication. *Interlanguage Studies Bulletin, 5,* 47-188.

Faerch, C. & Kasper, G. (1983). Procedural knowledge as a component of foreign language learners' communicative competence. In H. Boete & W. Herrlich (Eds.), *Kommunication im (Sprach-) Unterricht.* Utrecht.

Fillmore, L. W. (1976). *The second time around: Cognitive and social strategies in second language acquisition.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Fowler, H. W. & Fowler, F. G. (1977). *The concise Oxford dictionary of current English* (6th ed.). London: Oxford University Press.

Galloway, V. & Labarca, A. (1991). From student to learner: Style, process, and strategy. In D. W. Birchbichler (Ed.), *New perspectives and new directions in foreign language education* (pp. 111-158). Lincolnwood, IL: National Textbook Co. and American Council on the Teaching of Foreign Languages.

Green, J. (1991). Language learning strategies of Puerto Rican university students. Paper presented at the annual meeting of TESOL-Puerto Rico.

Horwitz, E. K. & Young, D. J. (Eds.). (1991). *Language anxiety: From theory and research to classroom implications.* Englewood Cliffs, NJ: Prentice-Hall.

James, B. G. (1984). *Business wargames.* Harmondsworth: Penguin.

Kellerman, E. Compensatory strategies in second language research: A critique, a revision and some (non-) implications for the classroom. In R. Phillipson, E. Kellerman, E. L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy research* (pp. 142- 161). Clevedon, UK: Multilingual Matters.

Krashen, S. & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning, 28,* 283-300.

Larsen-Freeman, D. & Long, M. (1991). *Introduction to second language acquisition research.* London: Longman.

Littlewood, W. (1979). Communicative performance in language development contexts. *International Review of Applied Linguistics, 18,* 123-138.

Marton, W. (1983). Second language acquisition tactics and language pedagogy. *System, 11*(3), 313-323.

Mayer, R. E. (1988). Learning strategies: An overview. In C. Weinstein, E. Goetz, & P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 11-22). New York: Academic Press.

McCombs, B. (1988). Motivational skills training: Combining metacognitive, cognitive, and affective learning strategies. In C. Weinstein, E. Goetz, & P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 141-170). New York: Academic Press.

Miller, G. A. , Galanter, E., & Pribram, K. H., (1960). *Plans and the structure of behavior.* New York: Holt.

Mischel, W. & Mishel, H. N. (1977). *Essentials of psychology.* New York: Random House.

Naiman, N., Frohlich, M., & Todesco, A. (1975). *The good language learner.* Research in Education Series, 7. Toronto: Ontario Institute for Studies in Education.

Nyikos, M. (1987). *The effect of color and imagery as mnemonic strategies on learning and retention of lexical items in German.* Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Omaggio, A. (1986). *Proficiency-oriented language teaching.* Boston: Heinle & Heinle.

O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. (1985a). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*(1), 21-46.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Russo, R., & Kupper, L. (1985b). Learning strategy applications with students of English as a second language. *TESOL Quarterly, 19*(3), 557-584.

Oxford, R. L. (1985). *A new taxonomy of second language learning strategies*. Washington, DC: Center for Applied Linguistics.

Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implications for strategy training. *System, 17*(2), 235-247.

Oxford, R. L. (1990a). Language learning strategies and beyond: A look at strategies in the context of styles. In S. S. Magnan (Ed.), *Shifting the instructional focus to the learner* (pp. 35-55). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Oxford, R. L. (1990b). *Language learning strategies: What every teacher should know*. New York: Newbury House/HarperCollins. (Now Boston: Heinle & Heinle).

Oxford, R. L. (1992/1993, Winter). Language learning strategies in a nutshell: Research update and ESL classroom implications. *TESOL Journal*.

Oxford, R. L. & Crookall, D. (1989). Research on language learning strategies: Methods, findings, and instructional issues. *Modern Language Journal, 73*(4), 404-419.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Politzer, R. L. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievement. *Studies in Second Language Acquisition, 6*, 54-65.

Poulisse, N. (1989). *The use of compensatory strategies by Dutch learners of English*. Unpublished doctoral dissertation, Catholic University of Nijmegen, The Netherlands.

Poulisse, N., Bongaerts, T., & Kellerman, E. (1984). On the use of compensatory strategies in second language performance. *Interlanguage Studies Bulletin, 8* (1), 70-105.

Pressley, M., Burkell, J., Cariglia-Bull, T., Lysynchuk, L., McGoldrick, J. A., Schneider, B., Snyder, B. L., Symons, S., & Woloshyn, V. E. (1990). *Cognitive strategy instruction that really improves children's academic performance*. Cambridge, MA: Brookline Books.

Rabinowitz, M. & Chi, M. T. (1987). An interactive model of strategic processing. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (pp. 83-102). Hillsdale, NJ: Erlbaum.

Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly, 9*(1), 41-51.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 11*(2), 118-131.

Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history, and typology. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 15-30). Englewood Cliffs, NJ: Prentice-Hall.

Scarcella, R. & Oxford, R. L. (1992). *The tapestry of language learning: The individual in the communicative classroom*. Boston: Heinle & Heinle.

Schmeck, R. (1988). Individual differences and learning strategies. In C. Weinstein, E. Goetz, & P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 171-191). New York: Academic Press.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Scovel, T. (1991). *The role of culture in second language pedagogy*. Position paper in Tapestry Program. Boston: Heinle & Heinle.

Seliger, H. W. (1980). Utterance planning and corrective behavior: Its function in the grammar construction process for second language learners. In H. Dechert & M. Raupach (Eds.), *Temporal variables of speech: Studies in honor of Frieda Goldman-Eisler*. The Hague: Mouton.

Seliger, H. W. (1982). Strategy and tactic in second language acquisition. Unpublished paper presented at the Summer School of Linguistics, Jyvastyla, Finland.

Seliger, H. W. (1983). Processing universals in second language acquisition. In F. R. Eckman, L. H. Bell, & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 36-47). Rowley, MA: Newbury House.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*(3).

Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.

Stern, H. H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review, 31*, 304-318.

Stern, H. H. (1983). *Fundamental concepts in language teaching*. Oxford: Oxford University Press.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. London: Cambridge University Press.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of human intelligence*. Cambridge University Press.

Stewner-Manzanares, G., Chamot, A. U., O'Malley, J. M., Kupper, L., & Russo, R. (1985). *Learning strategies in English as a second language instruction: A teacher's guide*. Rosslyn VA: InterAmerica Research Associates.

Sutter, W. (1989). *Strategies and styles*. Aalborg, Denmark: Danish Refugee Council.

Tang, B. T. (1990, Fall-Winter). From research on L2 learning strategies to classroom application. *TESOL-Gram, 17* (3 & 4), 13-16.

Tarone, E. (1977). Conscious communication strategies in interlanguage: A progress report. In H. D. Brown, C. A. Yorio, & R. Crymes (Eds.), *On TESOL '77: Teaching and learning ESL* (pp. 194-203). Washington: Teachers of English to Speakers of Other Languages.

Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning, 30*(2), 417-431.

Tarone, E. (1981). Decoding a primary language. The crucial role of strategic competence. Paper presented at the Conference of Interpretive Strategies in Language Learning, University of Lancaster, UK.

Tarone, E. (1983). Some thoughts on the notion of "communication strategy". In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 61-74). London: Longman.

Tarone, E., & Yule, G. (1989). *Focus on the language learner*. Oxford: Oxford University Press.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30*, 128-134.

Van Patten, B. & Lee, J. (1990). *Second language acquisition—foreign language learning*. Clevedon: Multilingual Matters.

Vann, R. & Abraham, R. (1989). Strategies of unsuccessful language learners. Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, San Francisco, CA.

Weinstein, C. E. & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315-327). New York: Macmillan.

Wenden, A. (1991). *Learner strategies for learner autonomy: Planning and implementing learner training for language learners*. Englewood Cliffs, NJ: Prentice-Hall.

Wenden, A. L. (1985). Learner strategies. *TESOL Newsletter, 19*(5), 1-7.

Wenden, A. L. (1986). What do second language learners know about their language learning? A second look at retrospective accounts. *Applied Linguistics, 7*(2), 186-205.

Wenden, A. L. (1987). A metacognitive study of L2 writers. Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, Miami, FL.

Wenden, A. & Rubin J. (Eds.), (1987). *Learner strategies for language learning*. Englewood Cliffs, NJ: Prentice-Hall.

Acknowledgment

The authors wish to thank the journal editor, Dr. Lidia Woytak, and the anonymous reviewers for their contributions to the revision of this paper.

Authors

REBECCA L. OXFORD, Associate Professor, Department of English, Curriculum & Instruction, The University of Alabama, 204 Graves Hall, Tuscaloosa, AL 35487. Specializations: language learning styles and strategies, language testing, language teaching methodology, psychological aspects of development.

ANDREW D. COHEN, Professor, English as a Second Language / Linguistics, University of Minnesota, 126 Klaeber Court, 320 - 16th Avenue Southeast, Minneapolis, MN 55455. Specializations: language learning and teaching, language testing, language teaching methodology, sociolinguistics.

Applied Language Learning
1992, Vol. 3, Nos. 1 & 2, pp. 37-63

Factors Relating to Learning a Foreign Language among High- and Low-Risk High School Students and Students with Learning Disabilities

Leonore Ganschow
Miami University

Richard Sparks
College of Mount St. Joseph

James Javorsky
Harold E. Fox Center, St. Joseph Mercy Hospital

Jon Patton
Miami University

This study examines results of a factor analysis on a battery of native language and foreign language aptitude tests administered to 80 high school students enrolled in first semester foreign language classes. Thirty-six students were identified as low-risk (LR) for learning a foreign language; twenty-nine as high-risk (HR); and fifteen as learning disabled (LD). The authors tested the validity of their hypothesis that there would be separable phonological, syntactic, and semantic factors, and that a phonological factor would have the most significant loading in the analysis. Three factors emerged from the principal component analysis and were defined as phonology/syntax (Factor 1); foreign language aptitude (Factor 2); and cognition/semantics (Factor 3). ANOVA procedures showed significant differences among the three groups on Factor 1. No differences were found between HR and LD groups on Factors 2 and 3, though the LR group differentiated itself from both groups. Results suggested that cognition/semantics may be distinguished from phonology, supporting recent research speculations on the modular and autonomous nature of the phonological code of language. High loadings on the Modern Language Aptitude Test suggest that this test may be a good predictor of possible foreign language learning problems when used in combination with other native language instruments.

Much of the foreign language (FL) literature in the past few years has focused on the importance of attitudes and motivation (Gardner, 1985, 1991), anxiety (Horwitz & Young, 1991), language learning strategies (Oxford, 1990), and other personality variables (Schumann, 1986) in the learning of a FL. However, recent attention has been given to physiological variables that might influence FL learning, in particular the impact that problems in ones development of native receptive and/or expressive language skills might have on ones performance in a FL (Skehan, 1986; Spolsky, 1989). In his examination of social, cultural, and other influences that affect FL learning, Spolsky identified 31 "necessary conditions . . . without which (second language) learning is impossible" (p. 14). Another 44 have been identified as "typically" involved in linguistic outcomes, but not required for them (Clark & O'Mara, 1991). Several of Spolsky's 31 necessary conditions involve physiological factors related to language: sound discrimination (#28); verbal memory (#29); and grammatical sensitivity (#30). Spolsky cites dyslexia, a language disability affecting ones ability to read, write and spell the language, as an example of a physiological condition that is likely to inhibit FL learning. He makes the Language 1/Language 2 link (hereafter referred to as *L1/L2*) explicit in Condition 22, the physiological normality condition, which states that ". . . any physiological or biological limitations that block the learning of a first language will similarly block the learning of a second language" (p. 89).

In the learning disabilities literature (which includes dyslexia), references to the difficulties that students with histories of reading/writing/spelling problems have in learning a FL in classroom settings have emerged only recently. (For a review of these studies, see Ganchow and Sparks [in press] and Sparks, Ganschow, and Javorsky [1992].) Increasing numbers of identified students with learning disabilities are now entering colleges and universities (Astin, Green, Korn, Schalit, & Berz, 1988) as a result of implementation of federal legislation protecting students in higher education. With this proliferation has come increasing frustration for faculty and staff who must grapple with the educational difficulties of these nontraditional students (Ganschow & Sparks, 1987; Mangrum & Strichart, 1988; Philips, Ganschow, & Anderson, 1991; Vogel, 1987). The requirement of FL study has been a major stumbling block for them (Freed, 1987; Ganschow & Sparks, 1987).

There is a small body of research that examines the relationship between L1 and L2. One line of research derives from studies on aptitude for learning a foreign language. For example, this link has been implied, though not explicitly stated, by Carroll and Pimsleur, who developed language aptitude batteries in the 1950s and 1960s.[1] Carroll and Sapon (1959), e.g., designed an artificial language in their *Modern*

Language Aptitude Test (MLAT) and identified four factors which in combination have been found to be predictive of aptitude for learning a FL in classroom contexts. These factors include phonetic coding ability (the ability to learn, recognize, and remember correspondences between the sounds of a language and their symbols), grammatical sensitivity (the ability to make judgments about the grammatical relations of sentences), rote memory (the ability to memorize and recall words in a new language), and inductive language learning (the ability to infer rules governing language). Pimsleur (1966) identified three factors in his *Language Aptitude Battery (LAB)*: verbal intelligence (familiarity with words in the mother tongue and ability to manipulate verbal materials); motivation (desire to learn the language); and auditory ability (sound discrimination, association of sounds with written symbols, sequencing of sounds/symbols, and memory for sounds/symbols).

In a L1/L2 linkage study Skehan (1986) found positive relationships between the development of first language proficiency and actual achievement in learning a FL in a longitudinal study of two cohorts of children born in England in the Bristol area. His findings showed two strong predictors of L2 aptitude: speed of development of first language and a cluster of factors, including L1 vocabulary, family class background, and literacy level of parents. Skehan proposed that second language learning aptitude is the FL equivalent of a first language learning capacity rather than just a residue of first language development. In other words, the elements that go into the learning of a first language are also present in the acquisition of a second language.

Further evidence for L1/L2 ties comes from research on the relationship between native language performance and FL aptitude among identified learning disabled (LD) and other at-risk populations. Research on college and high school at-risk populations by Sparks, Ganschow, and colleagues has shown that there are native language performance differences between good and poor FL learners.[2] In their *Linguistic Coding Deficit Hypothesis* (hereafter referred to as *LCDH*), the authors propose that difficulty with one's native language is a major factor in FL learning problems (Sparks & Ganschow, 1991; Sparks, Ganschow, & Pohlman, 1989). The hypothesis derives its name from research on reading disabilities by Vellutino and Scanlon (1986), who studied the role of the linguistic codes—phonology, syntax, and semantics—in good and poor readers. Phonology refers to the phonetic (sound/symbol) elements of a language; syntax involves understanding and then applying the grammatical and structural forms of a language system; and semantics involves word meanings and relationships between language and the concepts expressed in a language system.

Vellutino and Scanlon found that while poor readers have difficulty with both the phonological and syntactic aspects of language,

they do not have concomitant problems with semantics. The efficiency of the poor readers' phonological skills, in particular, is weak. Their research is consistent with an expanding body of empirical evidence in the reading disability literature showing that poor readers lack "phonemic awareness," or the knowledge that words are composed of phonemes and the ability to segment sounds within spoken words, often referred to as phoneme segmentation (Liberman, 1989; Stanovich, 1986a, 1988).[3] Good and poor readers do not appear to differ in their ability to use contextual skills, nor do they exhibit differences in semantic skills, e.g., language comprehension (Stanovich, 1986a, 1988). Rather, because of their phonological deficits, poor readers do not comprehend what they read, i.e., they cannot decode enough words to get the "gist" of the text. Poor readers also have difficulty with syntactic structures in both oral and written language (Vellutino & Scanlon, 1986; Vogel, 1987).

In their early FL case study analyses Sparks, Ganschow, and Pohlman (1989) studied the histories of 22 college students who had petitioned and received course waivers or permission to substitute for their institution's FL requirement. All of these students had failed or exhibited inordinate struggles in high school and/or college FL courses. These students had experienced histories of reading, spelling and writing problems. Results of the testing indicated that 68 percent of the students exhibited phonological deficits with or without concomitant syntactic and semantic deficits and 9 percent exhibited only a semantic deficit. All of the students had substantial difficulty with the language-based tasks on the MLAT. When each student's course failure history was compared to his/her linguistic coding profile, it was found that all but one student who had a phonological deficit had failed the college FL course in the first or second semester. The authors speculated that problems with the phonological code of language have an immediate and significant impact on performance in a FL course.

Over the years Sparks and Ganschow and their colleagues have evaluated large numbers of FL learners and have recently proposed five prototypes (Sparks & Ganschow, in press-b; Sparks, Ganschow, & Javorsky, 1992): (1) the "good" FL learner with strong phonology, syntax, and semantics; (2) the unsuccessful FL learner with weak phonology, weak or strong syntax, and strong semantics; (3) the unsuccessful FL learner with strong phonology, strong or weak syntax, and weak semantics; (4) the "garden variety" poor FL learner with weak phonology, weak syntax, and weak semantics; and (5) the unsuccessful FL learner with strong phonology, syntax, and semantics, but with affective differences such as low motivation, poor attitude, and/or high levels of anxiety. Sparks and his colleagues have found that the students with poor phonology and strong semantics (prototype 2) constitute the large majority of poor FL learners. In general, the poor

FL learners in their studies have shown remarkable similarities to the large majority of poor readers; that is, they exhibit difficulties with the phonological, and sometimes, syntactic, codes of language but do not have semantic problems.

Ganschow, Sparks and colleagues have conducted empirical studies at both the college and high school levels. At the college and university level Ganschow, Sparks, Javorsky, Pohlman, and Bishop-Marbury (1991) found significant differences in the phonological and FL aptitude (MLAT) skills of successful (A or B in two semesters of FL) and unsuccessful (failed high school or college FL courses and received course waiver or petition to substitute for the FL requirement) FL learners. No significant differences in IQ were found between the two groups. Also, no semantic differences were found on tasks of reading comprehension or receptive vocabulary. In another study, Ganschow, Sparks, Anderson, Javorsky and Skinner (1992) compared the native language and FL aptitude skills of high, moderate, and low anxious college FL learners identified using Horwitz's *Foreign Language Classroom Anxiety Scale* (Horwitz, 1991; Horwitz, Horwitz, & Cope, 1986). The three groups performed similarly on the ACT/SAT and a cognitive measure. However, significant group differences were found on measures of phonology, overall reading, oral language (listening and speaking), and FL aptitude. No significant differences were found among the three groups on measures of semantics (reading comprehension, vocabulary) or memory. In a third study (Ganschow & Sparks, 1991), an author-developed self report instrument composed of questions about students' academic learning and developmental histories was used to examine connections between native and FL learning. The authors found that questions relating to reading, spelling, writing, and grammar were the best discriminators for identifying students who were potentially at risk for FL learning problems. In another study on college students, Javorsky, Sparks, and Ganschow (1992) surveyed FL learners with and without learning disabilities on questions relating to their skills and attitudes about FL learning. Results of self-reports showed LD students perceived themselves as having fewer skills to meet the written and oral language demands of FL learning. No differences were found between the groups in their level of motivation to learn a FL, but LD students felt more anxious than students without learning disabilities about FL learning. The authors speculated that this higher level of anxiety was the consequence of the LD students' language learning problems.

At the secondary level, Sparks, Ganschow, Javorsky, Pohlman, and Patton (1992) compared low-risk and high-risk students enrolled in first-year FL courses. These low-risk (A or B grades in the first quarter of the FL) and high-risk (D or F in the first quarter of the FL) students were compared on measures of native language and FL aptitude. Significant differences were found between low- and high-risk learners

on all phonological and some syntactic measures and on the Short and Long Forms of the MLAT and all its subtests. No significant differences were found on any of the semantic measures. The authors suggested that the high-risk students had not been identified as potential problem learners before enrolling in the FL course because their native language syntactic and semantic skills were in the average range. It was only when phonological measures (e.g., spelling, phoneme segmentation, and pseudoword recognition) were administered that subtle deficits in the phonological skills of the high-risk group became apparent. In a related study (Sparks, Ganschow, Javorsky, Pohlman, & Patton, 1992b) the authors administered the same measures of native language and FL aptitude measures to a group identified as LD and compared them to the low- and high-risk groups. No significant differences were found between the high-risk and LD groups on most measures of phonology and syntax, as well as all subtests and the Short Form of the MLAT. Most important, perhaps, was that no significant differences were found among the three groups on any of the semantic measures (reading comprehension, vocabulary). These results support the authors' speculation that high-risk (non-LD) learners and students with identified LD share similar linguistic coding problems, manifested primarily in the phonological code.

In several articles, Sparks, Ganschow, and their colleagues have recommended the use of a *Multisensory Structured Language (MSL)* approach because it teaches the phonology and syntax of the FL (Myer, Ganschow, Sparks, & Kenneweg, 1989; Sparks, Ganschow, Kenneweg, & Miller, 1991). An MSL approach is derived from the Orton-Gillingham method, which traditionally has been used to teach reading-disabled (dyslexic) students to read, spell, and write their native language (Gillingham & Stillman, 1960). Supportive evidence for using this approach with at-risk FL learners who have phonological deficits is found in a study by Sparks, Ganschow, Pohlman, Skinner, and Artzer, (1992) in which two groups of at-risk FL learners enrolled in special self-contained sections of first-year Spanish were taught using an MSL approach. One of the groups was taught in both English and Spanish *(MSL/ES)*, the other in Spanish only *(MSL/S)*. A third group *(NO/MSL)* enrolled in a special, self-contained section of Spanish, was taught using traditional FL teaching methodologies combined with accommodations that have been used with LD students (e.g., frequent review, allowances for misspellings, emphasis on oral over written work, reduced homework assignments). The authors administered pre- and post-tests of native language phonology, syntax, and semantics, and FL aptitude (MLAT), and the pre- and post-test scores of the three groups were compared separately. Significant gains were made by the MSL/ES group on measures of native language phonology (pseudoword reading, phoneme segmentation, spelling, word recognition), vocabulary, and verbal memory, and the Short and Long Forms of the MLAT and all its subtests. The MSL/S group made significant gains on the Short and Long Forms of the MLAT. No significant gains were made by the NO/MSL group on any of the native language measures or the FL

aptitude test. The authors suggested that direct instruction in phonology may be a promising methodology for teaching a foreign language to at-risk students with phonological deficits.

In sum, results of these empirical studies all support the LCDH. That is, students with FL learning problems appear to have underlying and subtle native language learning problems and the great majority of poor FL learners exhibit difficulties primarily in the phonological (and sometimes syntactic) code of language. Problems with the semantic code do not appear to characterize most students with FL learning difficulties. The findings of these studies are in agreement with the reading research reviewed earlier suggesting that phonological, rather than semantic difficulties, are the primary problem of most poor readers. The LCDH also has support from factor analyses conducted by FL researchers in the 1950s and 1960s. The LCDH proposes language learning variables described as "necessary" by Spolsky and posits that without adequate native phonological, syntactic, and semantic language skills, students will have difficulty learning a FL in traditional classroom formats. The LCDH does not address Spolsky's "typical" or "graded" conditions, which would build upon the foundational (necessary) conditions. Although the LCDH acknowledges that these conditions exist, the authors of the LCDH speculate that for many FL learners the primary reason for FL learning success or failure lies in these initial "necessary" conditions (Sparks & Ganschow, 1991, in press-a).

Purpose of Study

In the present study the authors were interested in testing the validity of the LCDH in relation to an identified population of good and poor FL learners in high school. Of particular interest was whether separable phonological, syntactic, and semantic codes would emerge from a factor analysis of a test battery of native language and FL aptitude tests administered to these students. Because of recent studies showing the importance of phonology in both L1 and L2 learning, the authors speculated that tests measuring phonological skills would be the highest significant cluster and therefore the most important factor. Both Carroll (1962, 1973) and Pimsleur (1963) identified a similar factor, i.e, Carroll's phonetic coding and Pimsleur's auditory ability. However, neither specified that phonological knowledge is more important or more critical than other variables. In this study, an exploratory factor analysis procedure was used to test the validity of the LCDH.

Methodology

Subjects

The subjects were 80 high school students enrolled in first semester FL classes at eight suburban high schools. Six were public schools in Cincinnati and Columbus, Ohio. The other two were private

college preparatory schools in Cleveland and Baltimore. The FL courses represented in the study were Spanish, German, French, Latin, and Japanese. Among the population were 36 students identified as low-risk (LR) for learning a FL; 29 identified as high-risk (HR); and 15 identified as learning disabled (LD). The selection criteria for each group are reported here.

Low-risk (LR) and high-risk (HR). The participating LR and HR students were selected from a pool of 1200 students, who were administered an instrument designed by Ganschow and Sparks (1991) to screen students likely to be at risk for learning a FL. Scores deviating from an "ideal mean" are used to determine risk factor. Students scoring at least one standard deviation above the mean were identified as LR; those at least one standard deviation below the mean, as HR. LR and HR students were selected for further consideration. After selecting this initial pool, FL grades after the first 10-week grading period were used; those scoring an A or B were identified as LR; those scoring a D or F, as HR. Last, from the remaining pool, FL teachers were asked to recommend students they thought were LR and HR. The final selection pool resulted in 150 students being invited to participate. Of the 150 students, parental permission was obtained for 36 LR and 29 HR students, who formed the subject groups. In the LR group were 19 females and 17 males ranging in age from 14 years, 1 month to 17 years (mean age = 15 years, 5 months). In the HR group were 12 females and 17 males ranging in age from 14 years, 6 months to 17 years, 10 months (mean age = 15 years, 7 months). In order to maintain some degree of uniformity in cognitive ability and to utilize subjects having at least "average ability," the selected subjects were given a test of cognitive performance, the *Woodcock-Johnson Psychoeducational Battery: Brief Scale Cognitive Ability Cluster* (Woodcock & Johnson, 1978) (see description in *Figure 1*), and a cut-off of 90 (standard score) was established.[4] All of the selected subjects scored at or above 90. Subtests on this test were included as variables in the factor analysis.

Learning disabled (LD). LD subjects were identified by the local school district or private school using state guidelines for the diagnosis of LD. Fifteen students with LD were chosen to participate in the study. The students were tested using the aforementioned Brief Scale Cognitive Ability Cluster, and all of the selected subjects scored at or above 90.[5] In the LD group were 8 females and 7 males whose ages ranged from 14 years, 9 months to 17 years, 10 months (mean age = 16 years, 8 months).

Instruments

Figure 1 presents an alphabetical list and a description of the instruments used in the factor analysis procedure. The battery was selected to include tests measuring phonology (LAC; WRAT-R: Spelling Subtest; WJPB Reading Cluster: Letter-Word Identification and Word Attack subtests); syntax (WJPB Written Language: Dictation and Proofing subtests); semantics (Boston Naming Test; WJPB Reading Cluster: Passage Comprehension subtest); general cognitive ability (WJPB Brief Scale Cognitive Ability Cluster: Antonyms-Synonyms and Quantitative Concepts subtests); and FL aptitude MLAT Parts I, II, III, IV, and V.

FIGURE 1
Fourteen Tests/Subtests Used in the Factor Analysis

1. *Boston Naming Test* (Kaplan, Goodglass, & Weintraub, 1983) —tests ability to look at a series of pictures of graded difficulty and provide a vocabulary word (name) for each.
2. *Lindamood Auditory Conceptualization Test (LAC)* (Lindamood & Lindamood, 1979)—measures auditory perception and conceptualization of speech sounds.

3-7. *Modern Language Aptitude Test (MLAT)* (Carroll & Sapon, 1959)—tests foreign language aptitude in a simulated format and provides an indication of probable degree of success learning a foreign language. It consists of five subtests: *Part I: Number Learning; Part II: Phonetic Script; Part III: Spelling Clues; Part IV: Words in Sentences;* and *Part V: Paired Associates.*

8. *Wide Range Achievement Test (WRAT-R), Spelling Subtest* (Jastak & Wilkinson, 1984)—tests performance on writing single words from dictation.

9-14. *Woodcock-Johnson Psychoeducational Battery (WJPB)* (Woodcock & Johnson, 1978):

Brief Scale Cognitive Ability Cluster—is comprised of a combination of the *Antonyms-Synonyms* subtest, which tests knowledge of word meanings; and *Quantitative Concepts* subtest, which tests knowledge of quantitative concepts and vocabulary.

Written Language Cluster — has two subtests: *Dictation* subtest, which tests knowledge of punctuation and capitalization (8 percent of items), spelling (82 percent of items), and usage (10 percent of items); and *Proofing* subtest, which tests identification of grammatical errors including punctuation and capitalization (41 percent of items), spelling (24 percent of items), and usage (35 percent of items).

Reading Cluster — has three subtests: *Letter-Word Identification*, which tests ability to read isolated words; *Word Attack*, which tests ability to read pseudowords; and *Passage Comprehension*, which tests ability to determine a word that would be appropriate in the context of a passage (cloze test).

Procedure

Students were evaluated over a three-hour period. Tests were either individually administered or completed in small groups, according to test administration criteria. Testing was conducted by the authors, with assistance from special education majors trained by the authors prior to the study.[6]

Analysis of Data

Scores on the testing instruments were reported as standard scores, thus allowing adjustments for age differences and for making comparisons across age groups and among the various instruments.

In order to determine the dimensionality underlying the testing instruments, a Principal Components Factor Analysis with Varimax rotation was performed on the data. To allow for sampling variation, .9 was a minimum criterion for selecting the number of factors to be included (Jolliffe, 1986, p. 5). The exploratory factor analysis was performed in order to test the relationship among and between observed variables and their relationship to a set of unobserved factors or theoretical constructs (the *Linguistic Coding Deficit Hypothesis* or *LCDH*). This relationship appeared as a covariance, or a factor loading which has a range of 0 to 1.0. The higher the factor loading of an observed variable, the more significant the variable would be to the factor or the construct.

A standardized Scoring Coefficient was used to examine differences among the three populations in respect to the factors. A standardized scoring system based upon the factor loading of each variable was derived, which allowed the authors to use a One-Way Analysis of Variance (ANOVA) procedure to compare LR, HR, and LD groups on each factor.

Results

Three significant factors which were based upon the 14 tests emerged from the factor analysis. *Table 1* shows the three factors, their variables, their factor loadings, and the communality estimates of each of the variables. *Table 2* illustrates the variance accounted for by each factor. Factor 1 was defined as a Phonology/Syntax dimension; Factor 2, a FL Aptitude dimension; and Factor 3, a Cognition/Semantics dimension. The three extracted factors accounted for 73.1 percent of the total variance. Communality estimates showed that the Boston Naming Test, a vocabulary recall test, had the highest communality (.88), indicating that 88 percent of the variance of this test was explained by the three factors. The test having the lowest communality estimate (.51) was the MLAT: Part III .

TABLE 1

Factor Analysis and Communality Estimates on Test Battery *

Tests/Subtests	Factor 1 Phonology/ Syntax	Factor 2 FL Aptitude	Factor 3 Cognition/ Semantics	Communality Estimates
WJPB: Word Ident.	**.85**	.20	.29	.84
WRAT: Spelling	**.82**	.33	.22	.83
WJPB: Dictation	**.81**	.34	.28	.85
WJPB: Word Attack	**.80**	.29	.23	.78
WJPB: Proofing	**.76**	.30	.24	.73
MLAT: Part III	**.57**	.25	.32	.49
MLAT: Part V	.20	**.86**	.08	.79
MLAT: Part I	.26	**.73**	.30	.69
MLAT: Part IV	.41	**.60**	.37	.67
LAC:	.45	**.54**	.10	.51
MLAT: Part II	.44	**.52**	.43	.65
Boston Naming Test	.19	.12	**.91**	.88
Brief Scale WJPB: Cognitive	.44	.35	**.69**	.79
WJPB: Passage Comprehen.	.53	.29	**.60**	.73

* Boldfaced numbers indicate the factor on which the tests/subtests had the highest loadings.

TABLE 2

Variance Explained by Each Factor

	Description	Variance	Percent
Factor 1	Phonology/Syntax	4.78	34.1
Factor 2	FL Aptitude	2.94	21.0
Factor 3	Cognition/Semantics	2.52	18.0
Total variance accounted for		10.24/14	73.1%

Factor 1 obtained significant loadings (>.50) from 6 of the 14 measures and appeared to define the dimension *Phonology/Syntax*. Phonology/Syntax covers the linguistic codes of phonology and syntax and received heavy loadings from measures of spelling, word identification, word attack, and grammar.

Factor 2 obtained significant loadings from five measures relating to *FL Aptitude*. FL Aptitude received heavy loadings from measures of foreign and native language phonology, FL syntax and vocabulary, and variables measuring FL processing speed and rote memory.

Factor 3 obtained significant loadings on three measures and appeared to identify the dimension *Cognition/Semantics*. This factor was comprised of measures of vocabulary and cognitive ability and addressed the content of language, rather than the structure of the rules of language (syntax) or the sound/symbol system (phonology), i.e., Factor 1.

Table 3 presents the average standardized factor scores on each factor for LR, HR, and LD students. As the table shows, students with LD had the lowest average score on all three factors, especially Factor 1, the Phonology/Syntax dimension. Considerable variability in scores, as indicated by the large standard deviations, was noted for HR and LD groups for Factors 1 and 3, suggesting that there was large individual variation in performance among the classified groups. To demonstrate this variation visually and specifically, scattergrams depicting each subject's factor score were derived from the average standardized scoring coefficient. *Figures 2-4* show the scattergrams for Foreign Language Aptitude vs. Phonology/Syntax (*Figure 2*), Cognition/Semantics vs. Phonology/Syntax (*Figure 3*), and Cognition/Semantics vs. Foreign Language Aptitude (*Figure 4*). In general, the HR and LD groups (#'s 2 and 3) tended to fall closer to axes, whereas the LR group (#1) tended to cluster farther away. However, the frequency of statistical "outliers" among the HR and LD groups suggested that it was possible for an individual student to score low in one factor and high in another. This may further indicate that each factor appears to be measuring a unique and distinct dimension of language learning.

TABLE 3

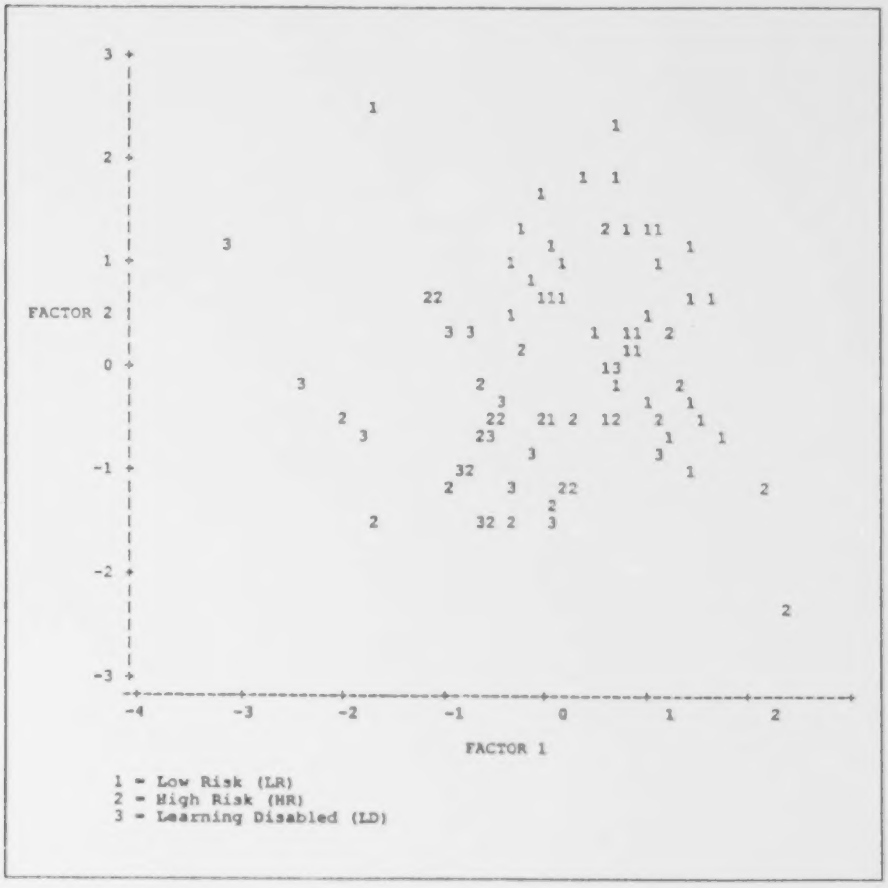
Average Standardized Score on Each Factor for Low-Risk, High-Risk and Learning Disabled FL Learners

Group	Factor 1		Factor 2		Factor 3	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
Low Risk (LR)	.50	.57	.58	.89	.35	.70
High-Risk (HR)	-.15	.93	-.56	.81	-.30	.98
Learning Disabled (LD)	-.95	1.21	-.37	.80	-.31	1.39

The scores here represent degree of deviation from the average standardized factor score for the entire group (N = 80).

FIGURE 2

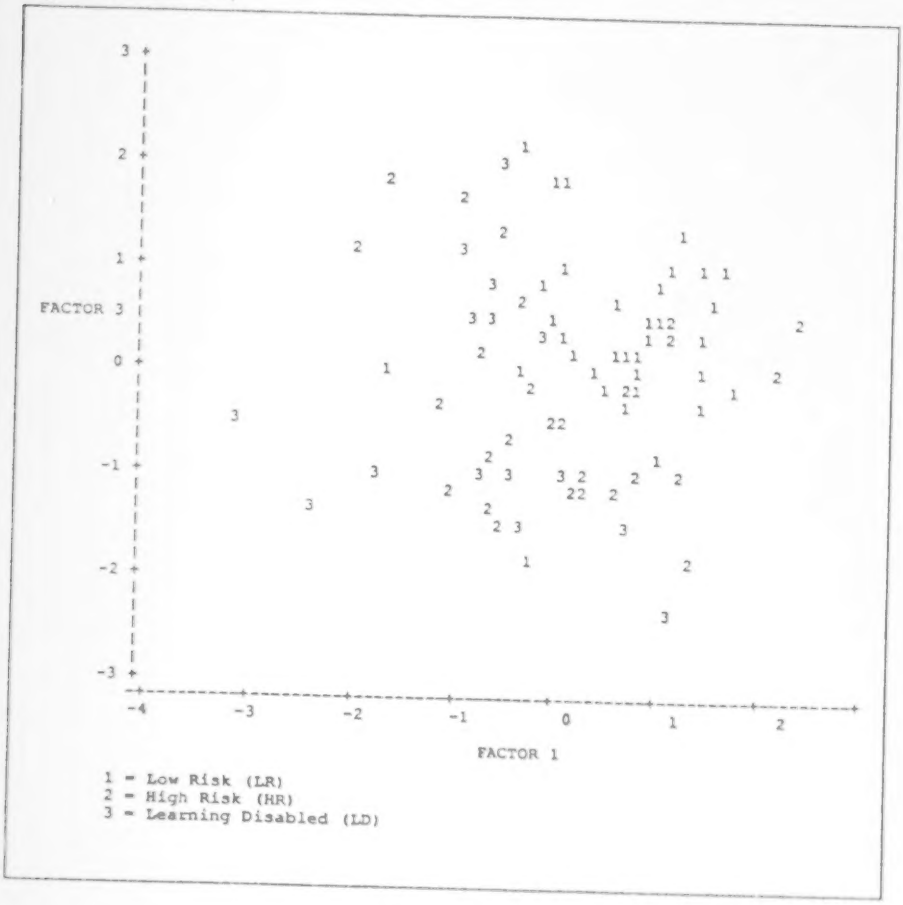
Scattergram of Foreign Language Aptitude (Factor 2) vs Phonology/Syntax (Factor 1)



BEST COPY AVAILABLE

FIGURE 3

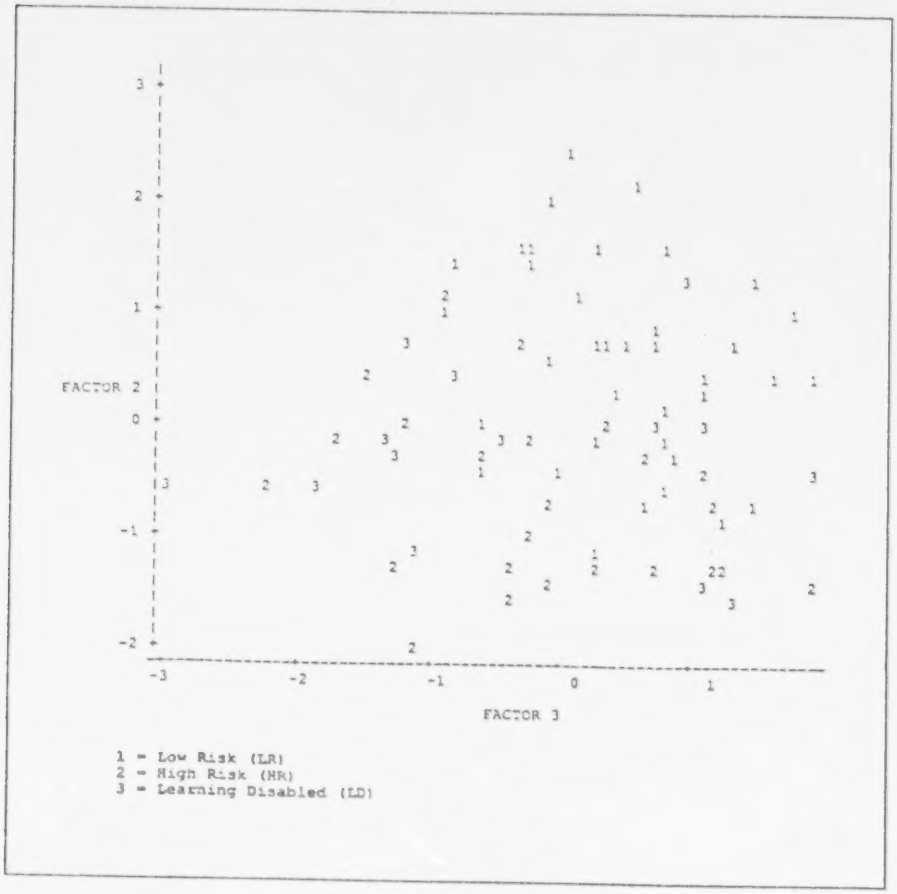
Scattergram of Cognition/Semantics (Factor 3) vs Phonology /Syntax (Factor 1)



BEST COPY AVAILABLE

FIGURE 4

Scattergram of Foreign Language Aptitude (Factor 2) vs Cognition/Semantics (Factor 3)



BEST COPY AVAILABLE

An ANOVA procedure, based on the standardized scoring system comparing HR, LR, and LD groups on each of the three factors, showed significant differences among the three groups on all three factors: Factor 1—$F(2,77) = 15.97$, $p = .0001$; Factor 2—$F(2,77) = 16.12$, $p = .0001$; and Factor 3—$F(2,77) = 4.66$, $p = .01$. A Scheffe procedure showed that on Factor 1 (Phonology/Syntax), the three groups differed significantly from each other. These differences were in descending order, with LD students scoring lowest, HR scoring second, and LR scoring highest. On Factor 2 (FL Aptitude), LR students, who had the highest score on this factor, were differentiated from HR and LD groups, but the HR and LD groups were not significantly different. On Factor 3 (Cognition/Semantics), LR students, who had the highest score on this factor, were differentiated from HR students but not from the LD group.

Discussion

Results of the factor analyses appear to support the LCDH, which proposes that there are separable language codes (phonological, syntactic, semantic). In this study three factors emerged from the factor analysis; they were identified as Phonology/Syntax, FL Aptitude, and Cognition/Semantics. From these findings, the authors infer that a factor they call Phonology/Syntax, as measured by the instruments in the battery, can be separated from general cognitive/semantic ability. This inference supports findings in the literature that cognitive processes employed in understanding and using the phonology (and syntax) of language may be substantively different from those employed in understanding and using semantics (vocabulary, comprehension). Recently, some theorists have postulated that there may be a specialized "module" for the phonology of language which is highly specific to the learning of the sound system of the language (Crain, 1989; Fodor, 1983; Shankweiler & Crain, 1986). This module is thought to be fast, automatic, and informationally encapsulated. It operates autonomously, rather than under the direction of higher-level language structures and does not interact with those structures. This module is also thought to be domain-specific, i.e., it operates only upon "speech" signals. By implication, then, it is possible to have difficulties in this component of language without immediate effects on the other language systems, i.e., the phonological module may not necessarily disrupt the global language processes needed for performance on vocabulary tasks, language comprehension, or intelligence tests (Stanovich, 1988).[7] The factor analysis conducted here supports the separability of phonology/syntax from these other cognitive processes.

The finding that phonology and syntax together formed the first factor initially suggested that there may be overlap between these two codes, as measured by the instruments in this study. However,

examination of the tests/subtests that comprise this factor (see *Table 1* and the description of them in *Figure 1*) shows that the syntactic elements being tested involve capitalization, usage, and spelling, with spelling comprising 82 percent of the items in the WJPB: Dictation subtest and 24 percent of the items in the WJPB: Proofing subtest. As spelling is considered a phonological skill (i.e., mapping sounds to letters), it might be argued that Factor 1 primarily measures skills related to phonology but only minimally measures skills considered to be syntactic or grammatical in nature.

The Phonology/Syntax dimension was the first, and therefore most important, of the three factors, accounting for 34.1 percent of the variance. LR and HR/LD FL learners in this study exhibited significant differences on Factor 1, which measured native language phonological and syntactic skills, and students at greatest risk were those with identified LD (*Table 3*). These findings on the relatively poorer performance by HR and LD students support research on reading (Liberman, 1989; Liberman & Shankweiler, 1985; Stanovich, 1986a, 1986b, 1988; Vellutino & Scanlon, 1986) and on foreign language (Ganschow, Sparks, Javorsky, Pohlman & Bishop-Marbury, 1991; Ganschow, Sparks, Anderson, Javorsky & Skinner, 1992; Sparks, Ganschow, Javorsky, Pohlman & Patton, 1992a, b; Sparks, Ganschow, & Pohlman, 1989) suggesting that students who have either native language or FL learning difficulties are likely to have particular problems with the phonological and grammatical (capitalization, usage) codes of language. FL researchers Carroll and Pimsleur both recognized the importance of these dimensions, reflected in their inclusion of measures of phonetic coding in their FL aptitude tests and Carroll and Sapon's (1959) inclusion of a measure of "grammatical sensitivity."

A specific factor for FL aptitude emerged from the study (Factor 2), suggesting its separability from native language skills. Given the authors' initial hypothesis that native and FL skills are both representative of underlying "language" skills, i.e., the L1/L2 linkage, an explanation should be offered to account for the emergence of this separate factor. An examination of the skills being measured by the MLAT may shed light on this point (*Figure 1*). Three of the five subtests (III, IV and V), i.e., the Short Form of the MLAT, are timed, and two of the subtests (I and V) are said to measure rote (verbal short-term) memory (Carroll & Sapon, 1959). Neither of these two skills (i.e., speed of performance and verbal short-term memory) was directly measured in any of the other tests/subtests in the battery. Yet these two processing skills have been found to distinguish good and poor readers (Mann & Liberman, 1984; Tunmer, 1989) and students with and without learning disabilities (Smith, 1991; Swanson, 1987). It is possible, then, that the FL "factor" as assessed by the four MLAT subtests that comprise Factor 2 may also be associated with such variables as short-term memory and

performance speed, both of which could be additional dimensions of general language ability. Skehan (1986) has argued for the differentiation of learner "subtypes" according to their performance on memory tasks. Likewise, he emphasizes differences in the time it takes students to learn a FL and suggests that both variables should be taken into account in FL instruction. While Skehan does not directly address speed of taking tests, he does discuss its importance, generally, in language learning.

Some explanation is necessary to account for the inclusion of MLAT: Part III (Spelling Clues) in Factor 1, and the Lindamood Auditory Conceptualization Test (LAC) in Factor 2. One might have expected the opposite to have occurred, i.e., the MLAT Spelling Clues subtest to have appeared in Factor 2, the Foreign Language Factor; and the LAC to have appeared in Factor 1, the Phonology/Syntax Factor. There are, however, several possible explanations for this anomaly. Carroll (1985) has indicated that though the MLAT: Part III is a "disguised" vocabulary test, it also measures phonetic coding ability because the vocabulary must be accessed through phonology (i.e., the student must look at a letter configuration such as *mblm* to "decode" the word *emblem* before choosing a word similar in meaning from five choices. Thus, success on the MLAT: Part III is at least partially dependent on phonology. With regard to the LAC, a phonemic awareness measure, a number of studies conducted by the present authors have shown differences in phonological skills between good and poor FL learners (see references by Ganschow, Sparks, and colleagues). As a result, the authors have speculated that phonemic awareness may be important in FL learning (as it has been shown to be in native language learning). For example, results of a recent predictive study conducted by three of the authors lend support to this speculation (Sparks, Ganschow, & Patton, 1992). In the study, phonemic awareness, measured by the LAC, was found to be one of the best four predictors of first-year final FL grades in a population of college-bound females enrolled in FL courses. This finding may help in part to explain the LAC's presence in the Foreign Language Factor.

Factor 3 (Cognition/Semantics) accounted for 18 percent of the variance and differentiated HR/LD groups from the LR group (*Table 3*). Results suggest that this dimension contributes to FL learning, though not to the same degree as phonology/syntax. The lower loading on the Cognition/Semantics factor might be accounted for by examining research on relationships between intelligence and other variables. For example, IQ tests have not been found to predict student achievement in a FL, although there is a modest correlation between intelligence and scores on FL aptitude measures (Carroll, 1985). Furthermore, FL aptitude tests (such as the MLAT) have been found to measure abilities that are distinctly different from those abilities measured by

intelligence tests (Wesche, Edwards, & Wells, 1982). Other studies also have shown that measures of general intelligence are relatively independent of language aptitude (Gardner & Lambert, 1965; Genesse, 1976). The evidence from FL researchers is in agreement with research conducted by native language researchers in reading, who have shown that although scores on intelligence tests tend to correlate more highly with vocabulary and other semantic measures, they do not correlate highly with phonological measures. In a recent paper, for example, Stanovich (1988) described evidence showing that early reading acquisition is predicted better by phonological measures than by standardized tests of intelligence.

One might hypothesize that Factor 3 involves what Fodor (1983) in his theory of modularity would call "central processes." Unlike modular processes, central processes operate slowly, have limited or no direct access to the information computed by modules, are more subject to voluntary control, and are unencapsulated. In Fodor's terms, modular processes are "blind" and "dumb"; central processes are "deliberate" and "smart."[8] On this dimension, unlike Factor 1, HR and LD groups appear to perform similarly. Yet, HR and LD groups differed significantly from the LR group, with HR/LD performing close to .5 standard deviation below the group mean and the LR group performing close to .5 standard deviation above the group mean (*Table 3*). The finding suggests that problems with the semantic code may play a less important role in FL learning problems, an inference that has been discussed elsewhere (Sparks & Ganschow, in press-a, Sparks, Ganschow, Javorsky, Pohlman & Patton, 1992a, b).[9]

Findings from observations of the scattergrams and the relatively larger standard deviations on Factors 1 and 3 by the HR/LD groups provide evidence for considerable variation in performance on these two factors. The possibility of distinct "prototypes" of FL learners, i.e., students with high phonology-syntax/low semantics, low semantics/high phonology-syntax, and "garden-variety" or overall poor performance on all codes, has been proposed elsewhere (Sparks & Ganschow, in press-b).

Implications

This study has several implications. One has to do with the diagnosis of at-risk FL learners; the second relates to instruction, and the third to implications for research on this population. In relation to diagnosis, the fact that the entire test battery accounted for a high proportion of the variance (73.1 percent) suggests that the battery may be potentially useful as a diagnostic indicator of students likely to be at risk for FL learning problems. However, the substantial standard deviations among the factors (*Table 3*) suggests that there is considerable variation among the subjects on these instruments,

especially among the HR and LD groups. Diagnostically, then, it would be helpful to look for subtypes of FL learning strengths and weaknesses among individuals by examining patterns of performance. A FL measure such as the MLAT, which has five subtests that measure relatively discrete skills (Carroll, 1962) should be considered when determining who is likely to be at risk for learning a FL. It is recommended that the whole battery (Long Form), rather than the Short Form (Subtests III, IV, and V), be used along with native language measures to diagnose a student for possible FL difficulties. Since Subtests III, IV, and V do not include any direct testing of phonology, the Long Form provides more diagnostic information than the Short Form. Also, it should be kept in mind that the latter three MLAT subtests are timed, which in itself may be a valuable indicator of language processing speed. Perhaps some of these students could learn a FL but are unable to do so in the given time frame. In support of FL aptitude testing, Skehan (1986) suggests that the field of aptitude testing has been sorely neglected and that it could play an important role in understanding individual learning differences. Skehan emphasizes the importance of being able to cope with auditory material and grammatical elements, which he refers to as "analytic" abilities (p. 210), and he recommends that tests measuring students' ability to use the phonological and grammatical aspects of the language should be employed. Because of the variation in learning strengths and weaknesses of at-risk FL learners, measures and methods for differentiating specific FL profiles or "prototypes" have been recommended (Ganschow & Sparks, in press-a, b; Sparks, Ganschow, & Javorsky, 1992).

A second implication involves instructional approaches for at-risk students who are attempting to learn a FL. One assumption underlying an instructional approach is that problems in L1 must be addressed either before or along with problems in L2. Demuth and Smith (1987), e.g., propose teaching the basic structure of native language prior to introducing the FL. A related approach addresses students' relative weakness in phonology/grammar by directly teaching these skills (Myer, Ganschow, Sparks & Kenneweg, 1989; Sparks, Ganschow, Kenneweg, & Miller, 1991; Wesche, 1981). Both FL and special educators have advocated using direct teaching of the phonology/syntax of language during the beginning stages. In FL instruction, e.g., Wesche (1981) specifically advocates "highly structured, incremental presentation of new material; frequent review of previously learned auditory material . . . and multisensory (including written) presentation and reinforcement of material" (p. 131) for students with phonological difficulties. In a multisensory structured language approach, students are taught the letters (graphemes) representing sounds (phonemes), and they are taught by simultaneously hearing, seeing, saying, and writing the FL sounds. Later, syntax is introduced in a similar multisensory fashion.

Preliminary empirical findings suggest that this method may be effective for at-risk learners (Sparks, Ganschow, Pohlman, Artzer, & Skinner, 1992).

In relation to the third implication, the need for further research, studies examining the role of short-term memory and the influence of pace/speed of learning should be investigated. Though there is substantial evidence to support short-term memory difficulties and the need for extensions of time for learning/testing among students with reading disabilities, there has been little research on these two skills in FL learning. However, recent research suggests that short-term phonological memory is valuable for long-term phonological learning, i.e., that phonological skills may play a role in the development of long-term memory skills, such as foreign language vocabulary acquisition (Papagno, Valentine, & Baddeley, 1991). Factor analyses using native language measures of short-term memory and timed measures along with the MLAT might also help to clarify whether the MLAT is tapping a specific aptitude for learning a FL or other cognitive processing skills.

This study provides additional support for the L1/L2 link clarified by Spolsky in his necessary conditions for learning a FL. It provides empirical support for the usefulness of the MLAT as one of the diagnostic indicators of potential difficulties with FL learning. It also demonstrates the importance of the phonological code (the awareness of the sound/symbol system of a given language, the ability to discriminate between sounds, and the meta-awareness that words are composed of sounds and how these sounds are put together to form words) in FL learning. In addition, the study found that phonology (and syntax) appear to be substantively different from the semantics (meaning elements) of language. Inferences from these findings have implications for both the identification and the instruction of students likely to be at risk for learning a FL.[10]

Notes

1. See Carroll (1962, 1973, 1974, 1981, 1985, 1990); Carroll and Sapon (1959); Pimsleur (1963, 1966); Pimsleur, Sundland, & McIntyre (1964).

2. See references by Ganschow and colleagues and Sparks and colleagues.

3. It is important to state that the authors' use of the term *phonology* does not refer primarily to ones ability to pronounce words in either the native or foreign language. It may include pronunciation but refers specifically to ones ability to learn sound (phoneme)/symbol (grapheme) correspondences, discriminate between sounds, and make explicit reports about sound segments in words. This latter skill, known as

phonemic awareness, involves "meta-awareness" of the phonemes within words (e.g., the word *last* has four phonemes, l-a-s-t).

4. This test has a Mean standard score of 100 and a standard deviation of 15. Thus, students with at least a 90 fall within the average range. Ninety represents the lower end of the Average range of intelligence (90-109). No restrictions were placed on how students could score on the IQ measure, although we did not expect to find many students with low average IQs (i.e., below 90) enrolled in FL courses. Of the 65 students for whom parent permission was obtained, no LR or HR student scored below 90 on the cognitive measure.

5. Parental permission was obtained for 20 students with LD. In order to maintain similarity in cognitive ability with LR and HR subjects, five students with LD who scored below 90 on the *WJPB: Brief Scale Cognitive Ability Cluster* were removed from the study, leaving 15 students in the LD group.

6. The authors would like to thank Vanessa Torbeck and Connie Yoho who assisted in administering the battery of tests.

7. It should be noted that long-term effects of an inefficient phonological module may include a less developed vocabulary, weaker listening skills, and poorer overall general knowledge (Stanovich, 1986a, 1986b, 1988) for a discussion of this issue.

8. See Fodor (1983) for an overview of the modularity hypothesis; also see Garfield (1991).

9. Is unlikely that the lower loading on Factor 3 is an artifact due to sample restrictions imposed by the score of 90 on the IQ test. First, IQ range restrictions were not used in screening the initial pool of 1200 students, selecting the pool of 150 students, or obtaining the final sample of 65 LR and HR students. Second, none of the 65 LR/HR students for whom parent permission was obtained scored below 90 on the IQ measure. Finally, the evidence from both foreign and native language educators indicates that cognitive measures do not contribute highly to the variance in FL learning.

10. The first two authors contributed equally in the preparation of this manuscript.

References

Astin, A., Green, K., Korn, W., Schalit, M., & Berz, M. E. (1988). *The American freshman: National norms for 1988.* Los Angeles: University of California.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training and research in education* (pp. 87-136). Pittsburgh, PA: University. of Pittsburgh Press.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of Psycholinguistics 2*, 5-14.

Carroll, J. B. (1974). Aptitude in second language learning. *Proceedings of the Fifth Symposium of the Canadian Association of Applied Linguistics*, 8-23.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-117). Rowley, MA: Newbury House.

Carroll, J. B. (1985). Second language abilities. In R. Sternberg (Ed.), *Human abilities: An information processing approach* (pp. 83-102). NY: W. H. Freeman & Co.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 1-29). Englewood Cliffs, NJ: Prentice-Hall.

Carroll, J. B. & Sapon, S. (1959). *Modern language aptitude test (MLAT) manual*. San Antonio, TX: Psychological Corp.

Clark, J. L. D., & O'Mara, F. E. (1991). Measurement and research implications of Spolsky's conditions for second language learning. *Applied Language Learning, 2* (1), 71-113.

Crain, S. (1989). Why poor readers misunderstand spoken sentences. In D. Shankweiler & I. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 133-165). Ann Arbor, MI: University. of Michigan Press.

Demuth, K. A., & Smith, N. B. (1987). The foreign language requirement: An alternative program. *Foreign Language Annals, 20*, 67-77.

Freed, B. F. (1987)). Exemptions from the foreign language requirement: A review of recent literature, problems, and policy. *ADFL Bulletin, 18*(2), 13-17.

Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Ganschow, L., & Sparks, R. (1987). The foreign language requirement. *Learning Disabilities Focus, 2*, 116-123.

Ganschow, L., & Sparks, R. (1991). A screening instrument for the identification of foreign language learning problems. *Foreign Language Annals, 24*, 383-398.

Ganschow, L., & Sparks, R. (in press). Foreign language learning disabilities: Issues, research, and teaching implications. In S. Vogel & P. Adelman (Eds.), *Success for college students with learning disabilities*. NY: Springer-Verlag.

Ganschow, L., & Sparks, R., Javorsky, J., Pohlman, J., & Bishop-Marbury, A. (1991). Identifying native language difficulties among foreign language learners in college: A "foreign"

language learning disability? *Journal of Learning Disabilities, 24*, 530-541.

Ganschow, L., & Sparks, R., Anderson, R. Javorsky, J., & Skinner, S. (1992). Differences in anxiety and language performance among high and low anxious college foreign language learners. Manuscript submitted for publication.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C. (1991). Second-language learning in adults: Correlates of proficiency. *Applied Language Learning, 2*(1), 1-28.

Gardner, R. C., & Lambert, W. (1965). Language aptitude, intelligence, and second language achievement. *Journal of Educational Psychology, 56*, 191-199.

Garfield, J. (1991). Introduction: Carving the mind at its joints. In J. Garfield (ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 1-13). Cambridge, MA: MIT Press.

Genesse, F. (1976). The role of intelligence in second language learning *Language Learning, 26*, 267-280.

Gillingham, A., & Stillman, B. (1960). *Remedial training for children with specific disability in reading, writing, and penmanship*. Cambridge, MA: Educators Publishing Service.

Horwitz, E. (1991). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 37-39). Englewood Cliffs, NJ: Prentice Hall.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70*, 125-132.

Horwitz, E., & Young, D. J. (Eds.) (1991). *Language anxiety: From theory and research to classroom implications*. Englewood Cliffs, NJ: Prentice Hall.

Jastak, S., & Wilkinson, G. (1984). *Wide range achievement test-revised (WRAT-R)*. Wilmington, DE: Jastak Associates.

Javorsky, J., Sparks, R., & Ganschow, L. (1992). Perceptions of college students with and without learning disabilities about foreign language courses. *Learning Disabilities: Research and Practice, 7*, 31-44.

Jolliffe, I. T. (1986). *Principal component analysis* (p. 95). NY: Springer-Verlag.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Liberman, I. (1989). The alphabetic principle and learning to read. In

D. Shankweiler, & I. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor, MI: University of Michigan Press.

Liberman, I., & Shankweiler, D. P. (1985). Phonology and problems of learning to read and write. *Remedial and Special Education, 6*, 8-17.

Lindamood, P., & Lindamood, C. (1979). *Lindamood Auditory Conceptualization Test.* Allen Park, TX: DLM Teaching Resources.

Mangrum, C., & Strichart, S. (1988). *College and the learning disabled student.* Orlando, FL: Grune & Stratton.

Mann, V. A., & Liberman, I. (1984). Phonological awareness and short-term memory. *Journal of Learning Disabilities, 17*, 592-599.

Myer, B., Ganschow, L., Sparks, R., & Kenneweg, S. (1989). Cracking the code: Helping students with specific learning disabilities. In D. McAlpine (Ed.), *Defining the essentials for the foreign language classroom* (pp. 112-120). Lincolnwood, IL: National Textbook Co.

Oxford, R. L. (1990). *Language learning strategies.* NY: Newbury House.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language, 30*, 331-347.

Philips, L., Ganschow, L., & Anderson, R. (1991). The college foreign language requirement: An action plan for alternatives. *NACADA (National Academic Advising Association) Journal, 11*, 51-56.

Pimsleur, P. (1963). Predicting success in high school foreign language courses. *Educational and Psychological Measurement, 2*, 349-357.

Pimsleur, P. (1966). *Pimsleur language aptitude battery and manual.* NY: Harcourt, Brace.

Pimsleur, P., Sundland, D. M., & McIntyre, R. (1964). Underachievement in foreign language learning. *International Review of Applied Linguistics, 2*, 113-150.

Schumann, J. H. (1986). Research on the Acculturation Model for second language acquisition. *Journal of Multilingual and Multicultural Development, 7*, 379-392.

Shankweiler, D., & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition, 24*, 139-168.

Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing, 3*, 188-221.

Smith, C. R. (1991). *Learning disabilities: The interaction of learner, task, and setting.* Boston: Allyn & Bacon.

Sparks, R., & Ganschow, L. (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *Modern Language Journal, 75*, 3-16.

Sparks, R., & Ganschow, L. (in press-a). Searching for the cognitive locus of foreign language learning problems: Linking first and second language learning. *Modern Language Journal*.

Sparks, R., & Ganschow, L. (in press-b). The impact of native language learning problems on foreign language learning: Case study illustrations of the linguistic coding deficit hypothesis. *Modern Language Journal*.

Sparks, R., Ganschow, L., & Javorsky, J. (1992). Diagnosing and accommodating the foreign language learning difficulties of college students with learning disabilities. *Learning Disabilities: Research and Practice, 7*(3), 150-160.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992a). Identifying native language deficits in high- and low-risk foreign language learners in high school. *Foreign Language Annals, 25*, 403-418.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992b). Test comparisons among students identified as high-risk, low-risk, and learning disabled in high school foreign language courses. *Modern Language Journal, 76*, 142-159.

Sparks, R., Ganschow, L., Kenneweg, S., & Miller, K. (1991). Use of an Orton-Gillingham approach to teach a foreign language to dyslexic/learning disabled students: Explicit teaching of phonology in a second language. *Annals of Dyslexia, 41*, 96-118.

Sparks, R., Ganschow, L., & Patton, J. (1992). Predictors of student performance in first-year high school foreign language courses (in preparation).

Sparks, R., Ganschow, L., & Pohlman, J. (1989). Linguistic coding deficits in foreign language learners. *Annals of Dyslexia, 39*, 179-195.

Sparks, R., Ganschow, L., Pohlman, J., Skinner, S., & Artzer, M. (1992). The effects of a multisensory, structured language approach on the native and foreign language aptitude skills of high-risk foreign language learners. *Annals of Dyslexia, 42*, 25-53.

Spolsky, B. (1989). *Conditions for second language learning*. Oxford: Oxford University Press.

Stanovich, K. (1986a). Explaining the variance in reading ability in terms of psychological processes: What have we learned? *Annals of Dyslexia, 36*, 67-96.

Stanovich, K. (1986b). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Stanovich, K. (1988). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia, 38,* 154-177.

Swanson, H. L. (Ed.) (1987). *Advances in learning and behavioral disabilities: Memory and learning disabilities.* Greenwich, CT: JAI Press.

Tunmer, W. E. (1989). The role of language-related factors in reading disability. In D. Shankweiler & I. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 91-132). Ann Arbor, MI: University of Michigan Press.

Vellutino, F., & Scanlon, D. (1986). Linguistic coding and metalinguistic awareness: Their relationship to verbal and code acquisition in poor and normal readers. In D. Yaden & S. Templeton (Eds.), *Metalinguistic awareness and beginning literacy* (pp. 115-141). Portsmouth, NH: Heineman.

Vogel, S. (1987). Issues and concerns in LD college programming. In D. Johnson & J. Blalock (Eds.), *Adults with learning disabilities* (pp. 239-275). NY: Grune & Stratton.

Wesche, M. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119-153). Rowley, MA: Newbury House.

Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics, 3,* 127-140.

Woodcock, R., & Johnson, M. B. (1978). *Woodcock-Johnson Psycho-Educational Battery (WJPB): Tests of Achievement, Cognitive Ability.* Allen Park, TX: DLM Teaching Resources.

Authors

LEONORE GANSCHOW, Associate Professor, Department of Educational Psychology, Miami Universxity, Oxford, OH 45056. Specializations: FL learning problems, language disabilities, reading diabilities.

RICHARD SPARKS, Associate Professor, Education Department, College of Mt. St. Joseph, Cincinnati, OH 45233. Specializations: FL learning problems, language disabilities, reading disabilities.

JAMES JAVORSKY, Psychoeducational Coordinator, Child and Adolescent Psychiatric Services, Harold E. Fox Center, St. Joseph Mercy Hospital, Pontiac, MI 48341. Specializations: language disabilities, reading diabilities, and behavioral/emotional disorders.

JON PATTON, Application Consultant & Adjunct Assistant Professor, Department of Decision Sciences, Miami University, Oxford, OH 45056 Specializations: applied mathematics and statistics.

Applied Language Learning
1992, Vol. 3, Nos. 1 & 2, pp. 65-85

The Utility of Objective (Computer) Measures of the Fluency of Speakers of English as a Second Language

Deborah Rekart *
Patricia Dunkel
The Pennsylvania State University

The speech of a sample of Chinese and Korean nonnative speakers of English (n=30) was evaluated by human raters of the Speaking Proficiency English Assessment Kit (SPEAK) for degree of fluency of the nonnative-English speech, and by a computerized silent-pause-detection program to determine whether: (1) the human-rated or the computer-rated indexes of fluency predicted subjects' SPEAK scores for overall comprehensibility in English; and (2) the correlation between the human-rated and computer-rated scores was significant. Findings indicated that human ratings of fluency correlated well with the human-rated overall comprehensibility score, but computer-assisted assessment did not. Intercorrelations among the human-rater evaluations were found, as were intercorrelations among the computer-rated scores; however, the two types of evaluations did not correlate with one another. The discrepancy between the two assessment methods led to the conclusion that humans use cues other than silent pauses to assess the fluency of nonnative speech even though the SPEAK guidelines stipulate that the human raters are to react to the amount of pausing contained in the nonnative stream of speech. The automatic pause-detection program reacted only to silent (unfilled) pauses in the speech signal. Speculations as to the cues human evaluators use are offered, and suggestions for objectifying fluency assessment of nonnative speech are presented.

* *The decision as to which researcher would be listed as first author was arrived at by chance with the flip of a coin.*

The number of international graduate students attending American graduate schools has increased steadily since the early 1980s. Many of these international students who are nonnative speakers of English serve as the teaching assistants of undergraduate native speakers of English in American universities. Although the international teaching assistants (ITAs) may possess excellent knowledge of their subject matter, they often lack strong skills in English, particularly spoken English. As a result, many of the undergraduates taught by ITAs report that they have difficulty understanding them. The undergraduates contend that many of the ITAs are neither comprehensible nor fluent in English, that is, that some of the ITAs lack oral proficiency in English (Anderson-Hsieh & Koehler, 1988). One element and measure of oral proficiency is fluency, defined by Beardsmore (1974) as the ability to sustain oral production, and by Fillmore (1979) as the ability to talk at length with few pauses in the stream of speech. These definitions focus on the ratio of the duration of speech to the amount of silence in the speech signal. The silence is a function of the amount of pause phenomena in the speech, including the amount of silent (unfilled) pauses, filled pauses, repetitions, false starts, parenthetical remarks, rate of speech, and correctness of articulation. According to Hieke (1980), the amount of pause phenomena contained in the stream of speech is an indicator of a nonnative speaker's level of fluency.

One standardized test of language proficiency that includes an assessment of oral fluency is the SPEAK test, a disclosed form of the Test of Spoken English (TSE). Both the SPEAK and TSE provide for assessment of a number of related language skills, including pronunciation, comprehensibility, grammar, and oral fluency. According to the guidelines of the SPEAK test, fluency in English as a second language (ESL) is to be measured in terms of the amount of pausing and fragmentation contained in the speech of the nonnative speaker (NNS). Fluency is evaluated on a scale from 0.0 to 0.4 (speech that is "so halting and fragmentary or has such a nonnative flow that intelligibility is virtually impossible") to 2.5 to 3.0 (speech that is "smooth and effortless, closely approximating that of a native speaker"). Speech containing "numerous nonnative pauses and/or a nonnative flow that interferes with intelligibility" would be rated between 0.5 and 1.4, whereas speech that shows "some nonnative pauses that do not interfere with intelligibility" would receive a fluency rating in the range of 1.5 to 2.4 (see *SPEAK*). The relative amount of pausing contained in nonnative speech influences not only the evaluator's judgment of the speaker's *fluency*, but also his or her judgment of the speaker's *overall comprehensibility*. The fact that the amount of pausing enters into the listener's judgment of the overall comprehensibility of the nonnative speaker is somewhat problematic since the same feature, pausing, is

used to describe two related yet distinct traits, fluency and comprehensibility.[1]

The overlap in the oral proficiency indicators of overall comprehensibility and fluency suggests that the constructs of comprehensibility and fluency have not been well conceptualized, or well operationalized by the test designers. It also suggests that the guidelines established to help explain and determine the fluency and comprehensibility scores in the TSE and SPEAK tests, commonly used metrics of oral proficiency for speakers of languages other than English, are imprecise and vague, at best. The imprecision and vagueness of the guidelines can lead to a lack of reliability in scoring. In addition, not only are the fluency and comprehensibility guidelines vague and overlapping, they are also extremely subjective. In Hieke's (1985) opinion, such vague and subjective guidelines are of little use because

> without recourse to the authority of a scientific framework, there can be no reliable norms, and such tests cannot be regarded as fair as long as they crucially hinge upon such prose statements to delineate levels while these are peppered with notions that cannot withstand close scrutiny (p. 137).

There is an additional problem with the SPEAK (and TSE) fluency guidelines. Although the guidelines for determining fluency and overall comprehensibility stipulate that the presence or absence of pauses in the speech signal are major determinants of perceived fluency, and should influence listener judgments of the comprehensibility of nonnative speech, they fail to provide unambiguous fluency (i. e., pause) criteria against which objective measures for each criterion can be derived. The guidelines also fail to distinguish specific quantitative and qualitative aspects of the several different kinds of pause phenomena (e. g., unfilled pauses not occurring at major syntactic boundaries) that distinguish nonfluent and incomprehensible speech from fluent and comprehensible nonnative speech. The authors of this report thought that the identification of the specific phenomenon of unfilled pauses might add to the knowledge base about the pause phenomena in question and might further help raters learn to assess the degree of nonfluency/fluency in the nonnative speech in more objective fashion. Such an objective measure could possibly result in a significant reduction in the amount of error involved in the listener-rater's assessment of the fluency of nonnative speech and, thereby, lead to increased accuracy of the ratings given.[2] The objective calculation of the amount of silent pausing in the speech of the English-speaking NNSs was made with the aid of a computerized automatic pause recognition program.

Computer-Assisted Analysis of Unfilled Pauses

Silent, or unfilled, pauses have been defined as silent intervals of some arbitrary duration (205 milliseconds, according to Goldman-Eisler [1968], or approximately 130 milliseconds, according to Hieke, Kowal, & O'Connell [1983]). Such a functional definition has provided the basis for some forms of machine analysis of periods of silence (i. e., pauses) as they occur in speech (Jensen, Ruder, & Harrington, 1972). Descriptions of automatic pause recognition devices are found in the literature as far back as Black, Tosi, Singh, and Takefuta (1966). Such devices respond to an absence of appreciable energy in the speech signal and yield data indicating the number, duration, and location of pauses, and the ratio of pause time to phonation time. One important advantage of pause detection by machine is that short pauses (200 milliseconds or less), which are not consciously perceived by listeners but which are detected by the physiology of the human ear, can be accurately tallied by computer (Tosi, 1974). In fact, O'Connell and Kowal (1983, p. 273) highlight the inability of human listeners to detect all the pauses in the flow of speech, claiming that

> the perceptual processes which enter into the experimenter's (or neutral judge's) estimates of time intervals in speech are profoundly affected by [the listener's] own native (and/or foreign) language habits. It is not simply a matter of not having a measuring instrument with sufficiently fine discrimination; it is a systematically biased instrument which perceives some unfilled pauses that physically do not exist and misses some that are longer than the ones it does register.

On the other hand, pausing also poses problems for the automatic (i. e., computer) pause recognizers. Jensen and his colleagues state that although silent interval duration is a major cue for the perception of pauses, listeners also recognize cues that the machine cannot. For example, several researchers (Jensen, Ruder, & Harrington, 1972; Martin, 1970; Mattingly, 1966) discuss the problems inherent in detecting cues associated with pausing, such as prepausal and postpausal effects.[3]

Since most prepausal and postpausal phenomena cannot be detected automatically by computer, only unfilled pauses over a specific threshold value (e. g., 200 milliseconds) set by the experimenter are detected automatically. In other words, pause recognizers are limited to the detection of one aspect of pause phenomena—silence; however, they are more accurate (i. e., more objective) in the detection of these silent, unfilled pauses than the human ear. In sum, with respect to the

detection and evaluation of pause phenomena, both human and computer assessments of fluency have strengths and weaknesses, and both are less than precise.

Investigations of First and Second Language Pause Phenomena

Like that of a native speaker, the speech of a NNS is characterized by unfilled pauses as well as other forms of pause phenomena (e. g., filled pauses and false starts). Deese (1978) posits that hesitation phenomena such as filled and unfilled pauses, as well as false starts, in nonnative speech reflect problems the NNSs have translating thoughts into words. Several researchers who have examined the issue of pause phenomena in the speech of nonnative speakers contend that the pausing in nonnative speech is related to the NNSs' level of semantic maturity and language proficiency. For example, Kowal and her colleagues (1975) found that the number of unfilled pauses in the reading aloud of NNSs decreased as proficiency in the L2 increased. O'Connell and his colleagues (1969) had previously found that variation in the semantic content of stories read by NNSs produced changes in the number and length of unfilled pauses used by German- and English-speaking adults when reading aloud. (Syntactic and other variations in the content and structure of the input were minimized in the experiment.)

In a follow-up study with a similar focus, O'Connell and Kowal (1972) examined pause phenomena occurring in the spoken English of native German-speaking and English-speaking subjects who read aloud and then recounted the stories read. In comparing the recounting by the native speakers of English and the native speakers of German, the researchers concluded that the pausing and hesitation in the speech of the nonnative speakers of English was a function of their relative proficiency in English. The results of the O'Connell and Kowal study suggested that lack of knowledge of L2 affects L2 fluency (i. e., fluency as measured by the amount of pause phenomena in the L2 discourse), and that amount of pausing in the speech signal might, therefore, constitute a measure of L2 oral proficiency. Just how sensitive a measure of proficiency it might be (or is), however, remains unclear and unvalidated.

A review of the small amount of research investigating L2 pause phenomena and their relation to nonnative fluency suggests that increasing levels of developed foreign language proficiency are associated with increasing levels of L2 fluency, and concomitantly with decreasing levels of pausing and hesitation phenomena. It might, therefore, be assumed that an assessment of foreign language proficiency should indeed include a measure of fluency, one aspect of which should comprise analysis of the ratio of unfilled pauses contained in the nonnative speech vis à vis the duration of the nonnative discourse. It was thought that the automatic pause-recognizer (i. e., a computerized

pause-recognition program) would be able to provide an objective measure of this ratio and could be used to gauge whether SPEAK raters react to the amount of silent pausing evident in the nonnative speech elicited for the SPEAK by NNSs of English. The present study was conducted to test this hypothesis.

Purpose of the Study

The basic purpose of this study was to investigate whether a computerized pause recognition program could be used to provide an objective measure of the oral fluency of NNSs of English, specifically speakers of Chinese and Korean. A second purpose was to determine whether a quantitative assessment of the amount of pausing in the speech detected by the computer correlated with the qualitative and subjective evaluation of the NNSs' fluency given by the human raters. That is, an attempt was made to determine what correlation, if any, existed between the quantitative amount of pauses recognized by the computer program and the qualitative judgment of the human raters of the fluency of Chinese and Korean speakers in two sections of the SPEAK test. Specifically, the subjective scores of subjects' fluency on two sections of the SPEAK, the picture description, and the syllabus explication, were correlated with two pause values, the mean number of pauses for the picture description and syllabus explication, obtained via an automatic pause recognizer. The power of the human-rated and computer-rated fluency measures to predict the nonnative speakers' overall comprehensibility evaluations on the test was also investigated by means of a stepwise multiple regression analysis.

The Research Questions

The basic research questions posed were the following:

(1) *Can a computerized pause-recognition program provide an accurate and objective measure of the amount of pausing found in the speech of nonnative speakers elicited in two sections of the SPEAK (the picture description and the syllabus explication exercises)?*

(2) *Does the amount of pausing detected by the computer program correlate with the subjective evaluation of human raters assessing (a) the fluency and (b) the overall comprehensibility of the nonnative speech?*

(3) *For the aspects of the SPEAK selected, which of the measures of fluency (i. e., the human-rated scores or the computer pause tallies) is the better predictor of nonnative speakers' overall comprehensibility scores?*

Method

Subjects

Thirty speakers representing two linguistic populations took part in the study: 9 speakers of Korean and 21 speakers of Chinese. The tapes of the speech of the Chinese and Korean speakers were selected for analysis for several reasons. The two groups were selected because more speakers from these groups than from other linguistic groups take the SPEAK at The Pennsylvania State University to ascertain whether they can assume duties as teaching assistants. Anderson-Hieke (1990) reports that there is a tendency for Chinese speakers to separate words rather than to link them in continuous speech and that "Korean ESL learners have some of the same suprasegmental problems as Chinese learners, although their problems have not been as well documented in the ESL literature" (p. 199).

Subjects' SPEAK overall comprehensibility scores ranged from 155 to 235. Overall Fluency scores ranged from 1.45 to 2.60. Fluency scores on the picture description (Section 4 of the SPEAK) ranged from 1.0 to 2.5. On the syllabus explication (Section 7 of the SPEAK), fluency scores ranged from 1.0 to 2.5.[4]

Materials and Procedures

Tape-recorded speech samples of the SPEAK test, which was administered to 30 international teaching assistants between 1986 and 1988 at The Pennsylvania State University, were examined. Samples of responses to Form I of the SPEAK were selected for analysis. In Form I, the picture description consists of a series of six pictures depicting a bicycle accident. The subjects were instructed to take one minute to prepare a story about the incidents illustrated in the pictures and then to spend one minute recounting the events. For the second task, the subjects received information that had to be explicated to an imaginary audience of students on the first day of class.

All the tapes were recorded in a group testing session in a language laboratory. The tapes were recorded on a Tandberg TAL 812 Computer-Controlled Advanced Learning Cassette Recorder. The speech samples were then copied with a Yamaha K-540 stereo cassette deck for subsequent analysis by computer. Out of a pool of 99 SPEAK tapes, a subsample of 30 Chinese and Korean subjects was selected for further analysis since they had all taken Form I of the SPEAK. The speech samples of the 30 subjects were digitized at a 12-kHz. sampling rate on an IBM/XT microcomputer with MetraByte analog-to-digital and digital-to-analog converters for the input/output and analysis of

speech. The computer was connected to an amplifier, stereo cassette deck, speakers, a digital filter, and a printer. The corpus was input and digitized using waveform editing routines, described by Drumright, Seikel and Wilcox (1987).

The corpus was input as a series of 10-second samples. Each consecutive sample overlapped with previous and succeeding samples by one waveform cycle in order to ensure the inclusion of all inter- and intra-sentential pause time. Approximately 15 soundwave files were created for each of the 3 subjects, for a total of 450 soundwave files.

Each soundwave file was analyzed for fluency using the automatic pause-detection program with flexible thresholds for noise level and minimal pause duration. The sound threshold for most files had to be set higher than the default value of 30 (i. e., 40-100) in order to detect pauses through tape noise. Minimal pause duration was set at 200 milliseconds to exclude the closure period for oral stop consonants /p, t, k, b, d, g/. Output from the program included a list of the number of individual pauses detected, with the duration and the starting time (in milliseconds) of each. In addition, data were provided regarding total pause time (in milliseconds), average pause length (in milliseconds), percent of total sample consisting of pause time, and pause/phonation ratio (comparison of duration of pause time or silence to duration of phonation or speech time). Although data regarding total pause time and the pause/phonation ratio were gathered, these data were not incorporated into the regression analysis since the small sample size ($n = 30$) limited the number of variables to five that could be used in the regression model. The automatic pause-detection information was verified by visual inspection of the waveform files screen-by-screen (53 milliseconds per screen) to ensure that each pause detected contained no phonation whatsoever (i. e., neither vocalization nor aspiration).

Statistical Analyses

Two types of statistical analyses were carried out: a stepwise multiple regression analysis, and Pearson product-moment analyses of the intercorrelations among all the human-rated and computer-rated (automatic pause-recognizer) scores.

Stepwise multiple regression analysis of the data was performed to identify whether the two computer-detected pause scores ($\#_p$, $\#_s$)[5] and/or the three human-rated fluency scores (FL_p, FL_s) and the human-rated aggregate score (OF) predicted the overall comprehensibility (OC)[6] scores of the NNSs on the SPEAK. Stepwise regression was also used to help determine which of these five indexes (the three human-rated fluency scores and two computer-rated silent pause scores) had the greatest predictive power relative to the SPEAK Overall Comprehensibility score, a basis for placement in the International Teaching Assistant (ITA) training classes at The Pennsylvania State University.[7] The data were analyzed with the main-frame SPSS[x] Batch system.

Results

Table 1 displays the correlations (R and R^2) and semipartial correlations (sr^2) among the variables, the standardized regression coefficients (ßs), the standard errors for the beta weights, the degrees of freedom, and the significant F ratios.

TABLE 1
Stepwise Regression of Fluency and Pause Scores for the Aggregate Score on the Test of Spoken English

OC	R	R^2	ß	SE ß	df	F	sr^2
OF	.81	.65	.81	9.64	1,27	51.22*	.65
FL_p	.86	.75	.36	6.81	2,26	39.56*	.10
FL_s	-	-	-	-	-	-	-
$\#_p$	-	-	-	-	-	-	-
$\#_s$	-	-	-	-	-	-	-

* $p < .01$.

If no value is displayed, no predictor proved significant. The symbols represent the following:

ß = *standardized regression coefficient*
SE ß = standard error of *beta*
sr^2 = semipartial correlation

OC = overall comprehensibility score SPEAK - human
OF = overall fluency score on SPEAK - human
FL_p = fluency score (picture description) - human
FL_s = fluency score (syllabus explication) - human
$\#_p$ = mean number of unfilled pauses (picture description) - computer
$\#_s$ = mean number of unfilled pauses (syllabus explication) - computer

According to Glass and Hopkins (1984), stepwise multiple regression computer programs show the increment added by each predicator (e. g., FLp) in the regression equation. In stepwise multiple regression, the best predictor is selected in Step 1 (OF), and a one-predictor regression equation is provided along with the correlation of rYY and statistics such as the standard error of estimate, etc.:

> In Step 2, the variable that would contribute the most additional relevant variance is selected and a two-predictor regression equation is selected and $RY_{1.2}$ is determined. The variable selected in Step 2 is the variable that has the highest corrrelation with Y when the previouly-entered independent variant is partialed out. Each successive step progresses in like manner—the next largest correlation with the criterion when all variables already included in the previous regression have been partialed out (p. 137).

BEST COPY AVAILABLE

Both the human-rated evaluation of overall fluency (OF) and the evaluation of fluency in the picture description section of the SPEAK (FL_p) proved to be major predictors of subjects' overall comprehensibility (OC) score (also a human-rated score). The stepwise regression analysis indicated that after Step 1, with overall fluency in the equation, $R^2 = .66$, $F(1,27) = 51.22$, $p < .001$. After Step 2, with the human-rated fluency of the picture description (FL_p), prediction of overall comprehensibility improved, $R^2 = .75$, $F(2,26) = 39.56$, $p < .001$. However, addition of the FL_p score resulted in a significant increment in R^2. The human-rated evaluation of the fluency for explication of the class schedule (FL_s) did not improve prediction of subjects' overall comprehensibility scores on the SPEAK. Neither of the computer-assisted tallies of pauses improved prediction of subjects' overall comprehensibility score on the SPEAK.

Pearson Product-moment intercorrelations were computed among human-rated scores (OC, OF, FL_p, FL_s) and the computer-rated indexes ($\#_p$, $\#_s$). *Table 2* presents the correlation coefficients.

TABLE 2
Pearson Product-Moment Correlations Among Human Ratings & Computer Pause Tabulations

	OC	OF	FL_p	FL_s	$\#_p$	$\#_s$
OC	1.00	.81**	.68**	.43*	−.15	−.30
OF		1.00	.51*	.54*	.00	−.06
FL_p			1.00	.14	−.13	−.32
FL_s				1.00	−.21	−.40
$\#_p$					1.00	.61*
$\#_s$						1.00

*$p < .01$; **$p < .001$ (1-tailed)

OC = overall comprehensibility score on SPEAK - human

OF = overall fluency score on SPEAK - human

FL_p = fluency score (picture description) - human

FL_s = fluency score (syllabus explication) - human

$\#_p$ = number of unfilled pauses (picture description) - computer

$\#_s$ = number of unfilled pauses (syllabus explication) - computer

BEST COPY AVAILABLE

Examination of these coefficients reveals a moderate to strong positive intercorrelation among the human-rated assessments. There were significant positive correlations between the OC score and (1) the OF score ($r = .81$, $p < .001$), and (2) the FL_p ($r = .68$, $p < .001$), and the FL_s ($r = .43$, $p < .01$). There was also a significant positive correlation between the OF and the FL_p ($r = .51$, $p < .01$) and the FL_s ($r = .54$, $p < .01$).

Intercorrelations were also found between the computer-rated scores. The number of pauses detected by the computer in the subjects' syllabus explication ($\#_s$) was moderately correlated ($r = .61$, $p < .01$) with the number of pauses detected by the computer in the subjects' picture descriptions ($\#_p$). The intercorrelation of the human-rated fluency scores for the picture (FL_p) and the syllabus explication (FL_s) was not significant. Rather interestingly, neither of the computer-rated tallies of pausing correlated with any of the human-rated indices. (See *Table 2.*)

DISCUSSION

Human-rated OF scores as well as the fluency score given for subjects' description of the pictures of the bicycle accident (FL_p) and the syllabus (FL_s) proved to be strong predictors of subjects' OC score (a score of the subjective estimates of the human raters). (In the picture description, the test taker has to compose orally without any vocabulary or discourse structure provided, unlike the syllabus explication condition.) It is interesting to note that the overall fluency scores of the human raters seems to explain, in large part, the overall comprehensibility score and suggests that there is a degree of consistency evident among the human rater scores (OF, OC, FL_p, FL_s). Apparently, the judgment made about overall comprehensibility by the human raters is influenced by the native listener's perception of the NNS's general level of fluency (i. e., her or his overall fluency) and the NNS's level of fluency when he or she is describing the series of pictures or the syllabus. There seems to be, in other words, a "halo effect" operating that is influencing the evaluator's judgment of the two traits (i. e., fluency and comprehensibility). Overall comprehensibility cannot be predicted by inclusion of computer-rated assessment of the number of silent pauses in the speech of the NNS. Obviously, the human-rated scores of overall fluency and fluency in describing the picture and the syllabus correspond well with the overall comprehensibility, another human-rated score, whereas the computerized tallies of pausing do not seem to be picking up on elements to which the human raters are responding, e. g., inappropriate prosody, incorrect lexical items, inaccurate syntax, or misarticulation).

Further research is needed to determine exactly what these elements are, and to what degree they influence the listener's judgments of the fluency and overall comprehensibility of the speech. Apparently, the amount of silent pausing in the speech is not the sole, nor possibly even a major, criterion of fluency. It seems likely that the humans are reacting to other components of linguistic competence, such as grammatical correctness, as well as to the amount, duration, and perhaps location (i. e., appropriate versus inappropriate location of a pause)[8] of the silent pauses in the speech signal. Although the SPEAK guidelines operationalize fluency in terms of the amount of silent pausing contained in the nonnative speech, the results of the present study indicate quite clearly that the evaluators of the nonnative speech were reacting to elements other than those stipulated in the guidelines. It is important that the guidelines given to evaluators, particularly to novice evaluators, articulate in clear and precise fashion exactly what constitutes the construct in question, in this case the fluency of nonnative speech. In their critique of the ACTFL Oral Proficiency Interview Guidelines, Bachman and Savignon (1986) suggest that the ACTFL Oral Proficiency Guidelines do not provide an unambiguous definition of "communicative language proficiency." The SPEAK Guidelines for the fluency component of the component scale might be faulted in similar fashion. As Bachman and Savignon suggest, testing guidelines need to include clear and unambiguous scale definitions and rankings of the illocutionary trait if we are ever to achieve a common metric of oral proficiency, one providing "criterion-referenced ratings that are comparable across a wide range of context and content areas" (1986, p. 388). The SPEAK scale definition of fluency does not clearly articulate the elements that constitute the ability to sustain oral production. The search to identify the relative contribution of these elements to the construct of fluency in L2 across levels of communicative language proficiency needs to be furthered, as does the interaction of fluency elements with other ingredients of oral proficiency.

The intercorrelations among the human-rated scores (OF, FL_p, FL_s) and the intercorrelations among the computerized tallies of pauses ($\#_p$, $\#_s$) are consistent and moderate to strong within each category (i.e., the human versus the computer category), but not across the two categories. This suggests a lack of congruency between the two types of assessments. It also suggests that the human evaluators may need to develop a better understanding of one important and discrete element of the *gestalt* of fluency: silent pausing in the speech signal. What may also be needed is for the human raters to understand the relationship that exists between the varying amounts (and kinds) of silent pausing in the speech sample and speakers' relative degrees of oral proficiency.

An objective estimate of the amount can be provided by a pause-recognition program. It would be interesting to observe whether and to what extent such information affects evaluations of fluency.

It is conceivable that in the future, artificial intelligence (AI) computer programs could be used to detect those elements in speech used by the human raters to make their judgments of fluency (e. g., the speaker's errors in syntax, prosody, and articulation, or the amount of silent pausing in the speech, and so forth). A computer program that yields a more comprehensive and accurate analysis of the speech signal than one based solely on automatic silent-pause detection is needed. Such a program would allow for more objective, accurate, and thorough assessment of a speaker's fluency than is presently possible. In the present study, the human evaluators responded to a host of indeterminate elements in the nonnative speech; the computer to only one. When we learn what elements the human evaluators are responding to when assessing fluency of nonnative English speech, we may be able to develop more sensitive fluency-detecting software that will speed up the tedious task of assessing oral fluency, one constituent member of the "set" of oral proficiency.

Researchers are just beginning to explore and to identify those elements that affect perceived degrees of fluency in the intuitive judgment of native-speaker (NS) listeners. In a microanalysis of the features of "highly fluent" and "highly nonfluent" speech of six Chinese speakers of English, Riggenbach (1991), for example, examined the influence of the following specific fluency/disfluency markers on NS listeners' perceptions of the Chinese speakers' relative degrees of fluency/disfluency: (1) hesitation phenomena (i. e., micropauses or silences of .2 seconds or less; hesitations of .3 or .4 seconds; unfilled pauses of .5 seconds or greater; and filled pauses); (2) repair phenomena (i. e., repetitions, restarts, false starts); (3) rate of speech (i. e., the number of words/semantic units per minute), and amount of speech (i. e., the total number of words/semantic units in raw frequencies, the percentage of NNS to NS speech and the total number of turns, including any contribution by either speaker, including "backchannels, unsuccessful attempts to claim turn, laughter particles, etc., as well as longer 'content' turns [p. 428])"; (4) interactive phenomena (i. e., continuers, echos, questions, and repair initiations); and (5) interactive features, (i. e., latches, overlaps, gaps between turns of .3 second or greater, and collaborative completions). Although Riggenbach reported that comparison of raters' fluency evaluations for the three "fluent" speakers and the three "nonfluent" speakers revealed few statistically significant differences in terms of some of the features listed above, she found that rate of speech and unfilled pauses—the latter a focus of the present study—contributed to NS listeners' judgments of NNSs' nonfluency.

Riggenbach concluded that

> Hesitation phenomena seem to be salient in determining fluency level, a fact intuited by designers of most widely used oral proficiency tests for second language learners. In this study, subjects could be broadly classified as less fluent or more fluent on the basis of the frequency of unfilled pauses alone. This possible trend—that more hesitations indicate lower fluency and fewer hesitations indicate higher fluency—would need to be validated with further studies of this type (and with much longer data bases). Also important, though not statistically significant in this study, is the placement of pauses, the degree of clustering or chunking of pauses with other disfluencies, pause type, and probable function of pauses (p. 438).

Riggenbach, who quantified pause phenomena previously discussed in Fillmore (1979) and Hieke (1980), calls for continued investigation of speakers "fluency profiles" in controlled investigations.

In an unpublished study that examined the influence of relative amounts of silent pausing in the English of native speakers of Chinese, researchers Rekart, Andrews, Tseng, Hung, and Gonzáles-Bueno (1991) focused on the perception of nonnative fluency by English-speaking listeners. The fluency of the Chinese-accented English speech sample which had been tape recorded was altered digitally, (1) by shortening the duration of four silent pauses in 25-ms steps, and (2) by removing the four pauses individually and in groups, and all together. Results revealed that 32 NS listeners perceived the NNS speech with longer pauses or with a greater number of pauses to be less fluent than speech with differences in pause amounts. Like Riggenbach, Rekart and her colleagues exhort that researchers conduct additional experimental studies to investigate the role of fluency features in listeners' perception of NNS fluency specifically, and oral proficiency in general. Such studies would help to pinpoint the components of fluency that most strongly influence native listener judgments, and the findings of such studies would, in turn, facilitate the development of superior fluency-detection software; the software could be incorporated into tests of oral proficiency in the future.

In a paper examining the psychological mechanisms underlying second language fluency, Schmidt (1992) suggests that researchers who attempt to pinpoint and to define the empirical correlates of perceived fluency should ask NNSs to introspect about their performance. Schmidt notes that the NNSs might not be able to introspect the macrostructure of automatic processes involved in fluency, that is, those that are "(a) fast and efficient, (b) effortless, (c) not limited by short-term memory capacity, (d) not under voluntary control, ((e) difficult to modify

or inhibit, and (f) unavailable to introspection" (p. 360). Learners may, however, be able to focus their attention on the controlled processes involved in fluent L2 speech, that is, those processes that are "(a) slow and inefficient, (b) effortful, (c) limited by capacity of short-term memory, (d) largely under subject control, (e) flexible, and (f) at least partly accessible to introspection" (p. 360). Schmidt hopes that asking learners to engage in introspection of this kind will "contribute to an understanding of the interplay of automatic and controlled processing in fluent and nonfluent L2 speech" (p. 379). Future investigations should heed Schmidt's exhortation.

CONCLUSION AND IMPLICATIONS

One of the designated criteria for evaluating fluency of speech in the SPEAK is the amount of pausing contained in nonnative speech; the findings of this study, however, demonstrate that the SPEAK evaluators are reacting to far more than unfilled pauses when evaluating the fluency and comprehensibility of the speech of NNSs. If the human evaluators had been responding to the amount of pausing, then there should have been strong correlations between the human-rated scores (FL_p and FL_s) and the computer-rated scores for the picture description ($\#_p$) and the syllabus description ($\#_s$). Such was not the case. There were intercorrelations among the human-rated scores, and there were intercorrelations among the computer-rated scores, but not across these two categories of scores. Apparently, the human raters were reacting to elements in the speech other than amount and duration of the pauses in the speech, but exactly what elements they were reacting to remains a mystery. Were they reacting to prepausal phenomena (e.g., phoneme lengthening), postpausal phenomena (e. g., partial devoicing of final consonants), hesitations, restructuring, fillers, grammatical errors, or a combination of these elements? We simply do not know. To answer such questions, perception experiments should be conducted in which listeners evaluate the fluency of speech in terms of pre- and post-pausal phenomena and other elements of linguistic competence.

To obtain a more objective measure of fluency, a computer-assisted assessment may be of use, but only if it can detect more pause phenomena than the number and duration of unfilled pauses. It must, in other words, permit detection of the location of the pauses as well as pre- and post-pausal phenomena, and other pause phenomena (e.g., filled pauses, hesitations, false starts, repeats, and parenthetical remarks). It is likely that some of these elements that are not automatically detected by the computer are playing major roles in influencing human raters' evaluation of the fluency and comprehensibility of nonnative speech. If such software were available, then more precise identification of the components of fluency would be possible, and more effective programs

for training SPEAK evaluators could be instituted. In addition, such information could be used in designing pronunciation programs aimed at helping ITAs improve their oral proficiency. At present, we rely on the evaluator's subjective and holistic impressions of fluency of the speech, and we rely on intuition and experience when designing pronunciation-remediation courses. It is hoped that more objective and precise assessment tools are forthcoming as advances in speech technology occur, and that more effective courses of instruction will be designed and implemented.

LIMITATIONS OF THE STUDY

Several limitations of the study need to be pointed out. The automatic pause-detection program detected only silent pauses; no other type of pause phenomena or pre- or post-pausal phenomena could be detected by the program.

An additional problem involves the subjective estimates of the NNSs' comprehensibility by the human raters. There was no independent assessment of what these comprehensibility ratings actually suggest about what was actually comprehended by listeners other than the SPEAK raters (such as student evaluations of the ITA subjects' overall comprehensibility):

> If [we] are going to talk about increasing the accuracy of assessments of comprehensibility, then it seems that the only way to determine if that has been achieved ultimately is to have some means of assessing comprehensibility in the target population, i. e., students of the ITAs (Dennis Gouran, personal communication, March 12, 1990).

The present study did not incorporate an independent assessment of the comprehensibility of the ITAs, although this construct was used as the dependent measure of analysis. Future research on the issue of accuracy of fluency and comprehensibility should include external measures of the trait under study.

The sample was small and comprised of native speakers of Chinese and Korean, so generalization to other populations of nonnative NNSs of English is naturally restricted. Furthermore, research on the oral fluency of other groups of nonnative English-speaking ITAs is sorely needed.

Acknowledgments

We would like to thank Dennis Gouran and Grant Henning for the substantive comments they offered reading a draft of the paper. We would also like to thank Rebecca McIlvaine, Cathy Andrews, Kim Steinhauer, and Ron Taylor for their research assistance.

Notes

1. The SPEAK scores for overall comprehensibility are reported on a scale that ranges from 0 to 300:

0 - 99	Overall comprehensibility too low in even the simplest type of speech.
100 - 149	Generally not comprehensible *because of frequent pauses and/or repharasing*, pronunciation errors, limited grasp of vocabulary, and lack of grammatical control.
150 - 199	Generally comprehensible with some errors in pronunciation, grammar, choice of vocabulary items, and *with some pauses or repharasing*.
200 - 249	Generally comprehensible with some errors in pronunciation, grammar, choice of vocabulary items, or *with pauses or occasional rephrasing*.
250 - 300	Completely comprehensible in normal speech, with occasional grammatical or pronunciation errors in very colloquial pharases.

2. Pennington (1991) discusses the use of the computer to achieve objective assessment of prosodic features across proficiency levels. She notes that graphic displays of certain prosodic features (e. g., epenthetic vowels), and the statistical calculations of these features by a system such as Kay Elemetrics *Visi-Pitch* allowed for objective comparison of the vowels in the speech of subjects at different levels of speaking proficiency. She also notes the *Visi-Pitch* can be used to describe and validate findings on the variation and acquisition of the prosodic features of intonation and stress. The graphic representations of the contours and the precise calculations and measurements can allow for a more exact description of pitch range and amplitude (intensity) variation (p. 134).

3. Klatt (1976, p. 1211) describes phoneme lengthening in the following fashion: "The syllable before the pause increases by about 60-200 milliseconds, with most of the durational increment restricted to the vowel and any postvocalic sonorants or consonants" (Oller, 1973: Klatt, 1975). Klatt offers an example of the phenomenon of phoneme lengthening by comparing the duration of the vowel [æ] in the word *bad*, in two sentences, *The cowboy and the badman led the carriage* ([æ] = 180 milliseconds) versus *The clumsy and the bad mangled the carriage* ([æ] = 300 milliseconds).

The prepausal effect of final obstruent devoicing refers to the phenomenon whereby voiced obstruents are partially voiceless when they occur before a pause (in utterance-final position). That is, voicing begins later in the sound than it does when the obstruents occur elsewhere in the utterance. For example, in the sentence, *The mice ran through the*

mazes, the second /z/ in /mezəz/ is partially devoiced at the end of the sentence, so that it sounds more like voiceless [s] than the first /z/ in the word. Prepausal effects may occur alone or in combination with postpausal effects. Lea (1973, p. 249) described one postpausal effect whereby sentence boundaries in his study were accompanied by large increases in fundamental frequency and substantial pauses. In 94 percent of all sentences in Lea's corpus, a decrease in fundamental frequency (FO) usually occurred at the end of each major constituent, and an increase in FO occurred near the beginning of the following constituent (the "fall-rise pattern"). For example, *Temperature records have* (FO down) *tumbled* (FO up) *in seven states.*" Cooper and Sorensen (1977) compared fundamental frequency contours at various syntactic boundaries including conjoined sentences, main clauses, embedded clauses, and phrases. They found that the fall-rise pattern was somewhat more pronounced for boundaries between two conjoined clauses such as the following: ([Anthony was surprised] and [Raymond became *upset*]], compared to other structures such as the embedded clause (*Anthony was surprised [Andrea became upset.]].*

4. Two pairs of trained NS teaching assistants of the Center for English as a Second Language at The Pennsylvania State University rated each of the tapes, thus providing the human evaluations of the international teaching assistants' speech. They received a total of 20 hours of training to judge the pronunciation, comprehensibility, and fluency of the NNSs. The training included use of the SPEAK training tapes, which provide samples of NNS speech that models a range of proficiency from 0 to 3 (see discusssion in the Introduction).

5. The computer-rated scores are represented by the following symbols:

$\#_p$ =	Mean number of unfilled pauses (picture description)
$\#_s$ =	Mean number of unfilled pauses (syllabus explication)

6. The human-rated scores are represented by the following symbols:

FL_p =	Human-rated for picture description (SPEAK section 4)
FL_s =	Human-rated score for syllabus explication (SPEAK section 7)
OC =	Human-rated score for overall comprehensibility (SPEAK)
OF =	Human-rated score for overall fluency (SPEAK)

7. Students achieving an Overall Comprehensibility score of less than 200 are required to take one semester of a Listening and Speaking class,

and two semesters of ITA training; students scoring between 200 and 229 in Overall Comprehensibility are required to take two semesters of ITA training; those scoring between 230 and 249 take one semester of ITA training. Those scoring above 249 are eligible to teach without taking the ITA training program.

8. An *appropriate* pause is defined as a pause which occurs at a syntactic (phrase, clause, or sentence) boundary. An *inappropriate* pause occurs elsewhere and includes (in this analysis) hesitations. An analysis of the appropriate and inappropriate pauses in a random sample (n = 9) segment of the corpus indicated that for human-rated fluency scores of 2.0 and 1.5, the cases in which the mean number of inappropriate pauses exceeded that of appropriate pauses were generally accompanied by high pause/phonation ratios, indicating extreme nonfluency of the speech; however, there was no apparent difference between the ratio of appropriate to inappropriate pauses or pause/phonation ratios in the higher-fluency, human-rated category of 2.0 as opposed to the lower-fluency, human-rated category of 1.5. That is, no pattern was evident in the sample of corpus drawn for the analysis of the relationship between appropriate and inappropriate pausing vis à vis the level of fluency as judged by the human raters. For human-rated fluency scores of 2.5 (n = 3), the speech was found to contain a higher ratio of appropriate to inappropriate pauses for 2 of the 3 speakers.

References

American Council on the Teaching of Foreign Languages. (1986). *ACTFL proficiency guidelines*. Hastings-on-Hudson, NY: ACTFL.

Anderson-Hsieh, J. (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes, 9*, 195-214.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking on native speaker comprehension. *Language Learning, 38*, 561-593.

Bachman, L., & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal, 70*, 380-397.

Beardsmore, H. B. (1974). Testing oral fluency. *International Review of Applied Linguistics, 12*, 317-326.

Black, J., Tosi O., Singh, S., & Takefuta, Y. (1966). A study of pauses in oral reading of one's native language and English. *Language and Speech, 9*, 237-241.

Cooper, W., Sorensen, J. (1977). *Fundamental frequency in sentence production*. New York: Springer.

Deese, J. (1978). Thoughts into speech. *American Scientist, 66*, 314-321.

Drumright, D., Seikel, J. A., & Wilcox, K. (1987). *Vocal jitter and shimmer measurements with a personal computer*. Paper presented at the Meeting of the Acoustical Society of America, Indianapolis, IN.

Fillmore, C. (1979). On fluency . In C. J. Fillmore (Ed.), *Individual differences in language ability and language behavior* (pp. 85-101). New York: Academic.

Glass, G. V. & Hopkins, K. (1984). *Statistical methods in education and psychology* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Goldman-Eisler, F. (1968). Psycholinguistics: Experiments in spontaneous speech. London: Academic.

Gouran, D. Personal communication. University Park, PA, 12 March, 1990.

Hieke, A. (1980). *Aspects of native and nonnative fluency skills*. Unpublished doctoral dissertation, University of Kansas.

Hieke, A. (1985). A componential approach to oral fluency evaluation. *Modern Language Journal, 69*, 135-142.

Hieke, A., Kowal, S., & O' Connell, D. C. (1983). The trouble with "articulatory" pauses. *Language and Speech, 26*, 203-214.

Jensen, P., Ruder, K., & Harrington, W. (1972). Pause adjustment mechanisms and measurement system (PAMMS). *Behavioral Research Methods and Instrumentation, 4*, 304-311.

Klatt, D. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics, 3*, 129-140.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208-1221.

Kowal, S., O' Connell, D., O' Brien, E., & Bryant, E. (1975). Temporal aspects of reading and speaking: Three experiments. *American Journal of Psychology, 88*, 549-569.

Lea, W. (1973). An approach to syntactic recognition without phonemics. *IEEE Transactions on Audio and Electroacoustics, AU-21*, 249-258.

Martin, J. B. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior, 9*, 75-78.

Mattingly, I. (1966). Synthesis by rule of prosodic features. *Language and Speech, 9*, 1-13.

O'Connell, D., & Kowal, S. (1983). Pausology. In W. A. Sedelow & S. Y. Sedelow (Eds.), *Computers in language research* (pp. 221-301). New York: Mouton.

O'Connell, D., & Kowal, S. (1972). Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research, 1*, 155-164.

O'Connell, D., Kowal, S., & Hormann, H. (1969). Semantic determinants of pauses. *Psychologische Forschung, 33*, 50-67.

Oller, J. (1973). The effect of position on speech segment duration in English. *Journal of the Acoustical Society of America, 54,* 123-147.

Pennington, M. (1991). Computer-assisted analysis of English dialect and interlanguage prosodics: Applications to research and training. In P. Dunkel (Ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 133-154). New York: Newbury House.

Rekart, D., Andrews, C., Tseng, R. M., Hung, Y.-C., & González-Bueno, M. (1991). *The importance of silent pauses in the perception of fluency.* Unpublished manuscript.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14,* 423-441.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14,* 357-385.

Speaking Proficiency English Assessment Kit. (1985). Princeton, NJ: ETS.

SPSS^X Batch System. University Park, PA: The Pennsylvania State University.

Test of Spoken English. (1985). Princeton, NJ: ETS.

Tosi, O. (1974). Pausometry: Measurement of low levels of acoustic energy. *Papers in Phonetics.* Tokyo: Phonetic Society of Japan, 129-144.

Authors

DEBORAH REKART, Department of Speech Communication, The Pennsylvania State University, University Park, PA 16802. Specializations: speech science and experimental linguistics.

PATRICIA DUNKEL, Department of Speech Communication, Center for English as a Second Language, The Pennsylvania State University, University Park, PA 16802. Specializations: second/foreign language listening comprehension; computer-adaptive testing of listening comprehension proficiency; computer-assisted language learning research.

From the Editor: A Note of Appreciation

The individuals listed below served as referees during the preparation of Volumes 1, 2, and 3 of *Applied Language Learning*. The editor and the editorial board express their gratitude for expert service to:

Kathleen M. Bailey
Christine M. Campbell
Francis Cartier
John L. D. Clark
Ray T. Clifford
Carolyn E. Corbett
Martha Herzog
Ruth L. Larimer
Betty Lou Leaver
John A. Lett, Jr.
Erika Malz
Monique Navelet
Rebecca L. Oxford
June K. Phillips
Victor M. Shaw
John Thain
Jean L. Turner
Leo van Lier

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editors encourage the submission of manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on $8\frac{1}{2} \times 11$ inch paper, double-spaced, with ample margins. Where feasible, manuscripts are preferred on disc (DOS or XEROX Viewpoint). Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

BEST COPY AVAILABLE

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

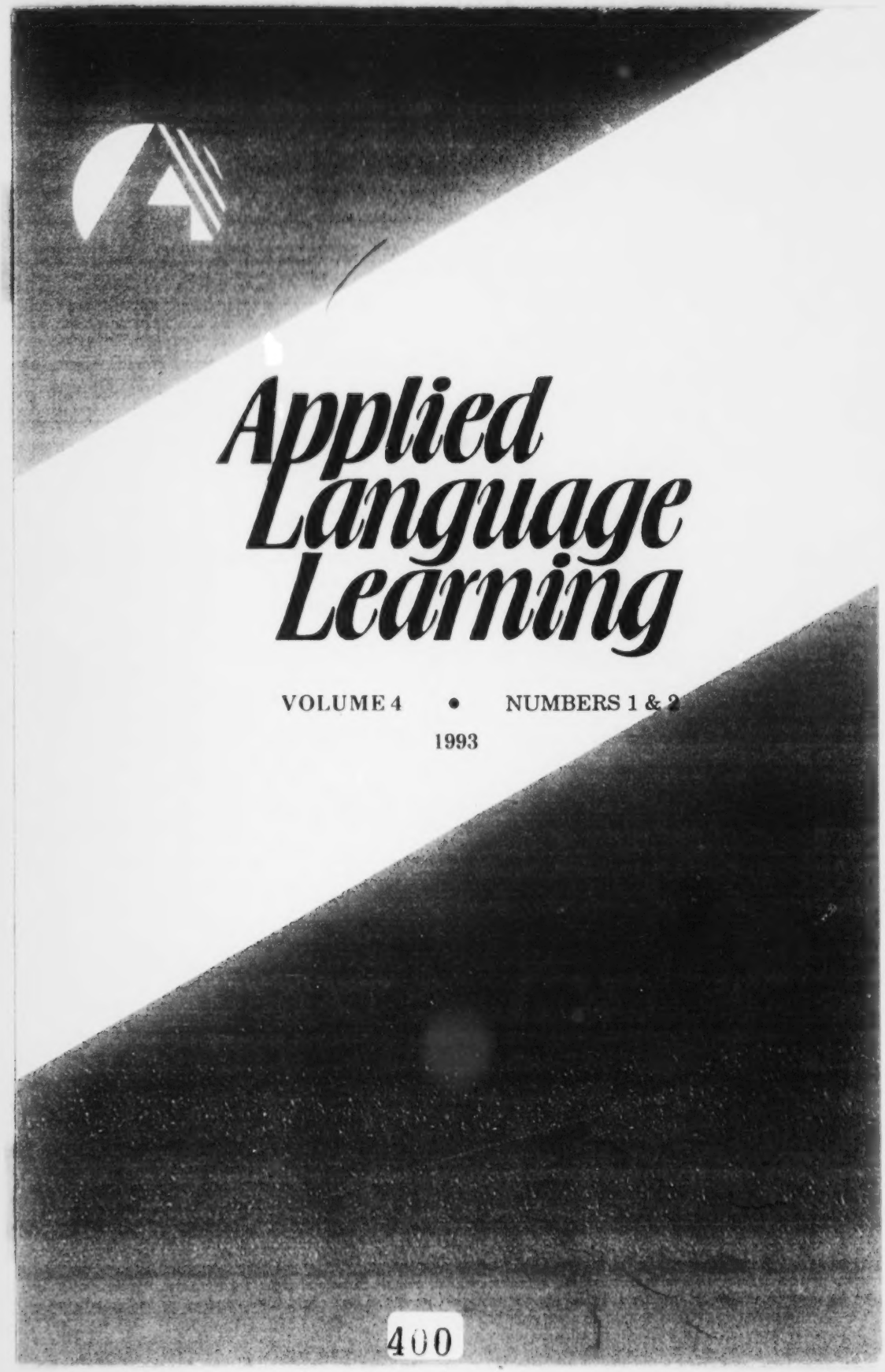
GPO 784-001/62163

PIN:070788-I/N-000

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

BEST COPY AVAILABLE

Professional Bulletin 65-92 1 & 2
United States Army
Approved for public release. Distribution is unlimited



Applied Language Learning

VOLUME 4 • NUMBERS 1 & 2

1993

400

Applied Language Learning

Volume 4 ❖ Numbers 1 & 2

1993

Defense Language Institute
Foreign Language Center

401

APPLIED LANGUAGE LEARNING

PB 65 - 93 - 1 & 2
US ISSN 1041-679

The mission of *Applied Language Learning (ALL)* is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Professional Bulletin 65-93, *Applied Language Learning* published semi-annually by the Defense Language Institute, Foreign Language Center, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

Official:

Milton H. Hamilton

MILTON H. HAMILTON
Administrative Assistant to the Secretary of the Army
04245

GORDON R. SULLIVAN
General, United States Army
Chief of Staff

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to DLIFLC.

Applied Language Learning is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Book rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster — Send change-of-address information to: *Applied Language Learning*, Office of the Provost, Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. Telephone: (408) 647-5638 or DDN: 878-5638.

APPLIED LANGUAGE LEARNING

PB 65 - 93 - 1 & 2

1993 • Volume 4 • Numbers 1 & 2

Lidia Woytak
Editor

Editorial Board

Kathleen M. Bailey	Professor, TESOL/TFL Program, Monterey Institute of International Studies
David P. Benseler	Professor, Case Western Reserve University
John L. D. Clark	Dean, Program Evaluation, Research, and Testing, DLI
Ray T. Clifford	Provost, DLI
Martha Herzog	Dean, Curriculum and Instruction, DLI
Betty Lou Leaver	Dean, Slavic School, DLI
John A. Lett, Jr.	Director, Research, DLI
June K. Phillips	Director, Tennessee Foreign Language Institute
Gary D. Walter	Librarian, DLI

DLI Staff

Editorial Coordinators, Sabine Atwell, Christa Rutsche
Production Editor, Jack Boyd
Associate Editor, Mike McGuire
Cover Design, Barney Inada
Graphics, Leonard Kodl
Layout Consultant, Annemarie Frazier

APPLIED LANGUAGE LEARNING

PB 65 - 93 - 1 & 2

1993 • Volume 4 • Numbers 1 & 2

CONTENTS

1 From Learning-Centeredness to Learner-Centeredness
David Nunan

19 Proficiency and Performance in Language Testing
James Child, Ray T. Clifford, and Pardee Lowe, Jr.

55 Incidental Vocabulary Acquisition in French as a Foreign Language
Beatrice Dupuy and Stephen D. Krashen

65 Instructional Implications of Gender Differences in Second/Foreign Language (L2) Learning Styles and Strategies
Rebecca L. Oxford

95 Sheltered Initiation Language Learning
Zev bar-Lev

131 Calendar of Events

136 Information for Contributors

Applied Language Learning
1993, Vol. 4, Nos. 1 & 2, pp. 1-18

From Learning-Centeredness to Learner-Centeredness *

David Nunan
Macquarie University

Two closely-related concepts are explored in this article: learner-centeredness and learning-centeredness. Both concepts have aroused interest and controversy in recent years, and each is invested with a certain mystique. In the first part of the article, I shall attempt to provide an interpretive gloss on these terms, followed by a selective review of research since the publication of The Learner-Centered Curriculum (Nunan, 1988). In the final part of the article, I would like to present practical ideas for experimenting with learner-centeredness in foreign language teaching.

What is Learner-Centered Teaching?

Before talking about research and practice, it is important to distinguish between learner-centeredness and learning-centeredness. To begin with learner-centereedness, the basic idea behind the concept is simple, although not simplistic. A learner-centered curriculum will contain components similar to those in traditional curricula. The principal difference is that in a learner-centered curriculum, key decisions about what will be taught, when it will be taught, and how it will be assessed, will be made with reference to the learner. Information about learners and, where feasible, from learners, will be used to answer key questions of what, how, when, and how well. Other themes which have emerged in writings on learner-centeredness include an emphasis on active involvement of learners in communicating in the classroom (Scarcella & Oxford, 1992); a focus on the use of authentic materials (Cathcart & Vaughn, 1993); and incorporating into the curriculum learning-how-to-learn goals alongside language goals (Ellis & Sinclair, 1989).

My own interest in seeing things from the learner's point of view developed when I realized that I could not do the learning for my learners—that in the final analysis, they would have to do their own learning, and that the best thing I could do was to help them find ways of

* I wish to thank Kathi Bailey for her insightful comments on an earlier version of this paper.

© 1993, David Nunan

doing their own learning. Supporting this notion is the realization of the complex and indirect relationship between instruction and learning. When I began as a language teacher, I thought, perhaps with the arrogance of the young, that my learners ought to learn what I taught. It was a shock to discover how naive was the idea that learners learn what teachers teach. It is a fact of life that learners do not learn what teachers teach in the simplistic one-to-one way implied by many curricular specifications and assessment tools.

However, it is a serious error to leap from an acceptance of the notion that learners must ultimately do their own learning to embracing the doctrine of learner autonomy. While I believe that learner autonomy (that is, the freedom, responsibility and power of learners to choose what they want to learn and how they want to learn) is an end toward which all teachers and learners ought to work, it is a mistake to assume that learners come into our classrooms with some kind of natural endowment to choose both wisely and well. In fact there is substantial evidence to suggest that many learners are not so naturally endowed. This is where the notion of learning-centeredness makes its appearance. A learning-centered classroom features dual complementary aims. While one set of aims is focused on language content, the other is focused on the learning process.

Learners are therefore systematically educated in the skills and knowledge they will need in order to make informed choices about what they want to learn and how they want to learn. Rather than assuming that the learner comes to the learning arrangement fortified with critical learning skills, the sensitive teacher accepts that many learners will only begin to develop such skills in the course of instruction.

Learner-centeredness, then, is not an all-or-nothing concept. It is simply not the case that classrooms are either teacher-centered or learner-centered. It is a relative matter. As I have already implied, it is also not the case that a learner-centered classroom is one in which the teacher hands over power, responsibility, and control to the students from day one. I have found that it is usually well into a course before learners are in a position to make informed choices about what they want to learn and how they want to learn, and it is not uncommon that learners are in such a position only at the end of the course.

Let us look briefly at the different stages in planning, implementing, and evaluating a language course, and compare what might happen in a learning-centered classroom, which is designed to develop learning skills in the student, with what, for want of a better term, we might call the institution-centered classroom.

At the syllabus planning stage in the institution-centered classroom, it is the institution or the teacher who makes all the decisions about what will be taught and when it will be taught. These decisions will be made with little or no reference to the actual or potential communicative needs of the learner. In a learner-centered classroom, on

the other hand, the selection and sequencing of content will be made with reference to the sorts of uses to which the learner will want to put the language outside the classroom.

In selecting learning experiences in a learning-centered classroom, the teacher will introduce a range of learning activities and tasks. There will also be an attempt to identify the learning style preferences of the learner, using these as the starting point in making methodological selections. There is a further element which distinguishes a learning-centered from an institution-centered classroom. In the former, the learners will be encouraged to reflect on their learning experiences, and to evaluate the opportunities made available to them in class. In this way they learn not only about the target language, but about the learning process itself.

In terms of assessment and evaluation, the learning-centered classroom places great store on training learners in techniques of self-assessment, ongoing monitoring, and self-evaluation. Once again, the learning-centered teacher does not assume that learners have these skills at the beginning of the learning process or that all learners will appreciate the potential value of self-monitoring to the development of communication skills in the target language. However, during the course of instruction, they will be provided with opportunities for engaging in self-monitoring activities and using these as a way of developing their language skills as well as their sensitivity to the learning process. (Brindley [1989] provides an excellent introduction to learner assessment and self-assessment, along with practical ideas for involving learners in their own monitoring and assessment.)

Recent Research: Implications for the Learner-Centered Classroom

In the preceding section, I outlined some of the conceptual underpinnings to learner- and learning-centeredness. However, the questions which need to be confronted are: Is there any evidence that learner- and learning-centered teaching actually works? Is there any evidence that the active engagement of the learner in the learning process has any effect on learning outcomes? Do learners have views on what they want to learn and how they want to learn it?

Some years ago, I carried out a comparative study into the learning preferences of teachers and learners in the Australian Adult Migrant Education Program (Nunan, 1986). When I compared the preferences of learners and teachers in relation to selected learning tasks and activities, I found stark contrasts and dramatic mismatches. The results of this study are summarized in *Table 1* (see p. 5). From the table, it can be seen that there are mismatches between teachers and learners on all but one of the items. (In this item, students and teachers

agreed that conversation practice was a very high priority.) In all other cases, there were mismatches between the teaching preferences of the teachers and the learning preferences of the students. For example, students gave a low rating for pair work, while teachers gave this item a very high rating. The same was the case with student self-discovery of errors.) Now I am not suggesting that student views should be acceded to in all cases. However, I would argue that at the very least, teachers should find out what their students think and feel about what they want to learn and how they want to learn.

In 1988, Ken Willing, one of my colleagues at the National Centre for English Language Teaching and Research, carried out a large scale study into the learning styles and learning strategy preferences of adult immigrant learners of English as a second language in Australia. With 517 learners, Willing had a substantial data base. Using a questionnaire and interview, Willing investigated possible learning-style differences which could be attributed to range of learner biographical variables such as ethnic background, age, level of education, time in the target country, and speaking proficiency level. The study came up with certain surprising findings. In the first instance, learners did have views on the learning process, and were capable of articulating these. Perhaps the most surprising finding was that none of the biographical variables correlated significantly with any of the learning preferences:

> None of the learning differences as related to personal variables were of a magnitude to permit a blanket generalization about the learning preferences of a particular biographical sub-group. Thus, any statement to the effect that 'Chinese are X,' or 'South Americans prefer Y,' or 'Younger learners like Z,' or 'High School graduates prefer Q' is certain to be inaccurate. The most important single finding of the study was that for any given learning issue, the typical spectrum of opinions on that issue was represented, in virtually the same ratios, within any biographical subgroup. (Willing, 1988)

In Europe, Dam and Gabrielsen (1988) investigated the extent to which young learners were capable of making decisions about the content and processes of their own learning. Beginning in the early 1980s, and continuing for a 6-year period, they investigated whether

Table 1
Teacher / Student Mismatches (Nunan, 1986)

ACTIVITY	STUDENT	TEACHER
explanation to class	very high	high
conversation practice	very high	very high
error correction	very high	low
vocabulary development	very high	high
student self-discovery of errors	low	very high
using pictures, film, video	low	low medium
pair work	low	very high
language games	very low	low

eleven-year-old learners of English as a foreign language in Denmark were capable of being involved in planning, organizing, managing, and evaluating their own learning. They also wanted to know how such a collaborative process could be maintained over a 6-year period, and what effect such a collaborative approach would have on the development of communicative competence in their learners. They found that learners, regardless of their aptitude or ability, were capable of a positive and productive involvement in selecting their own content and learning procedures. Furthermore, learners were also positive in accepting responsibility for their own learning. Interestingly, the researchers found that difficulties reported by teachers had less to do with learner resistance than with the problems of redefining teacher roles.

In a very different context, one of my own graduate students in Australia, Assinder (1991), came up with a similar result when her students were given an opportunity to determine what content was to be learned and how the content was to be dealt with in class. Assinder developed a peer-teaching peer-learning model, in which students prepared video materials to present to one another. Her own observations, as well as student evaluations, confirmed the success of this experiment in student-initiated teaching and learning. Assinder reported increased participation and responsibilty, increased accuracy, and sustained motivation. The learner-oriented nature of the experiment is evident in Assinder's conclusion:

> I believe that the goal of "teaching each other" was a factor of paramount importance. Being asked to present something to another group gave a clear reason for the work, called for greater responsibilty to one's own group, and led to increased motivation and greatly improved accuracy. The success of each group's presentation was measured by the response and feedback of the other group; thus there was a measure of built-in evaluation and a test of how much had been learned. Being an expert on a topic noticeably increased self-esteem, and getting more confident week by week gave [the learners] a feeling of genuine progress. (Assinder, 1991, p. 228)

Widdows and Voller (1991) sought the views of university students in Japan. They report that, contrary to popular wisdom, these learners have definite views on what they want to learn and how they want to learn. These views conflict with the official curriculum of the university. In their report, they state that:

> The most important result of this survey is the dichotomy between what students want to learn and experience in university English classes, and what they are actually taught there. . . . Students do not like classes in which they sit passively, reading or translating. They do not like classes where the teacher controls everything. They do not like reading English literature much, even when they are literature majors. Thus it is clear that the great majority of university English classes are failing to satisfy learner needs in any way. Radical changes in the content of courses, and especially in the types of courses that are offered, and the systematic retraining of EFL teachers in learner-centered classroom procedures are steps that must be taken, if teachers and administrators are seriously interested in addressing their students' needs.

A study by Kumaravadivelu (1991) focuses on the multiple perspectives on reality in the classroom. His research demonstrates the mismatch between teacher intention and learner interpretation, and underscores the naivete of assuming that learners will learn what it is that teachers intend for them to learn. His study identifies ten potential sources of mismatch between teacher intention and learner interpretation and shows how, if we are aware of these, differing interpretations need not be negative, but "if identified and properly handled can give learners an opportunity to negotiate further in order to tease out a problem in their own way." (Kumaravadivelu, 1991, p. 196).

Heath (1992) investigated what happens when ESL learners in the United States are collaboratively involved in selecting content, reflecting on learning tasks, and involved in their own self-evaluation. She used a simple but nonetheless ingenious device for collecting her data. She wrote to the high school students of an associate, asking them to collect data on their uses of oral and written language. These students were all in Basic English classes. They had major deficiencies in their written language skills and some of them were considered to be "mentally deficient." Heath asked these students

> to work together as a company of ethnographers, collecting, interpreting, and building a bank of information about language in their worlds. They had access to knowledge I wanted, and the only way I could get that knowledge was for them to write to me. They collected field notes, wrote interpretations of patterns they discovered as they discussed their field notes, and they answered the

> questions I raised about their data collection and their interpretations. (Heath, 1992. p. 42).

As a result of their involvement in this process, the students learned that communication is negotiation, and they got to reflect on the important relationships between socialization, language and thought. In substantive terms, all students moved out of the basic English in regular English classes, and two moved into honors English. As Heath reports, "Accomplishments were real and meaningful for these students."

In addition, Heath was able to conclude that:

> If the students learned that communication is negotiation ...what did the teacher and researcher learn? [We] gradually recognized that the best way to assess students' progress was to ask them to analyze their own written work. They had recorded, described, and interpreted patterns of oral and written language in their homes and the work settings of their parents, as well as their own classrooms. Why couldn't they look at their earlier letters and compare them with later letters? . . . The evaluative process, usually imposed by the teacher, became one in which students helped determine how one letter might be "better" than another, and how far they had come between September and May in their writing. (Heath, 1992, p. 43).

Another study employing a deceptively simple methodology, but from which some fascinating results emerged is Slimani (1992). Slimani sought to determine what learners actually learn from classroom interaction. Her question was as follows: "What is it that individual learners claim to have learned from interactive classroom events?" (This claimed learning she terms "uptake.") She further asks, "What is it that happens in the lesson that can account for this uptake?" Her learners were a group of Algerian learners of English as a foreign language who were preparing to undertake engineering studies in English. Slimani found that topics initiated in the classroom by the learners were much more likely to be nominated as having been learned than those nominated by the teacher. In other words, when learners had an opportunity to contribute to the content of the lesson, that was the content which learners would claim to have learned.

In a major study in Singapore, Lim (1992) sought evidence for a correlation between classroom processes and learning outcomes. In this study involving 1,600 junior college students, she found that the

frequency or quantity of learner participation related significantly to qualitative aspects of learner participation such as the range of speech acts and the control of conversational management techniques (e.g. turn maintenance and topic initiation). Learner participation in class also related significantly to improvements in language proficiency as measured by proficiency tests. Finally, there was a high correlation between activation of language outside the classroom and participation in class.

Lim's results are mirrored in an investigation by Hall (1991). Hall wanted to know how two-way information gap tasks, in which students have to share different kinds of information in order to reach a teacher-planned outcome, contribute to the acquisition of mathematical vocabulary. Subjects for the study were 11- to 13-year-old students of mixed proficiency. There were three matched groups of eight students in the experiment. The experimental group used split information tasks in a pair work learning arrangement. The first control group used the same content, but completed the tasks as an individual worksheet without the benefit of interaction. The second control group received teacher-fronted instruction. All students were pre-tested, underwent a 4-week instructional period, and were then retested on four different measures: multiple choice comprehension, sentence anomaly detection, free recall writing, and an oral interview.

While the post-test scores showed that all groups in Hall's study benefited from the task-based learning, the group which worked interactively on the information gap tasks did so to a significantly greater extent than the two control groups. Hall detected a strong correlation between the generative use of vocabulary items and the students' success in manipulating the items during attainment tasks. He concluded that "it would thus appear that spoken output in the context of split information tasks is an important factor in aiding the acquisition of target vocabulary items." (Hall, 1991)

These studies are summarized in *Table 2*.

Table 2
Recent Research with Implications for the Learner-Centered Classroom

STUDY	LEARNERS	OUTCOMES
Nunan (1986)	adult ESL learners in Australia	Learners have definite views on what they want to learn and how they want to learn. There are often dramatic mismatches between the views of the teacher and the learners.
Willing (1988)	adult ESL learners in Australia	Learners have different learning strategy preferences which seem to relate to personality rather than to biographical variables, such as ethnicity or education.
Dam & Gabrielsen (1988)	11-year-old EFL learners in Denmark	Learners are able to take responsibility for planning, organizing, managing, and evaluating their own learning.
Kumaravadivelu (1991)	adult ESL learners in the United States	There are identifiable sources of misunderstanding between teachers and learners in the classroom.

Widdows & Voller (1991)	university students in Japan	Students want to be involved in the selection of language content and learning process. Major mismatches exist between the desires of the student and the mandated curriculum.
Slimani (1991)	Algerian college-level EFL learners	Classroom topic nominated by learners were much more likely to become "uptake" than topics determined by the teacher.
Lim (1992)	junior college students in Singapore	Learner participation in class relates significantly to improvements in language proficiency.
Heath (1992)	high school students in the United States	Learning is enhanced when students are activity-involved in selecting content, learning tasks, and evaluation.
Hall (1991)	ESL school students in New Zealand	Spoken output in group work is a significant factor in the acquisition of vocabulary.

These studies represent only a sample of all those having implications for learner/learning-centeredness that have been carried out in recent years. However, the overall thrust of the research is clear.

From the studies, we see that:

- Learners are different and learn in different ways.
- There may be as many different interpretations of what is going on in the classroom as there are learners in the class.
- Learning is enhanced when learners are actively involved in communicating with each other and with the teacher in the target language.
- Learning is enhanced when learners are given opportunities to select content and learning tasks and also when they are provided with opportunities to evaluate their own progress.

Some Practical Suggestions

In 1989, I was asked to go to Japan to run some seminars on learner-centered teaching. I was rather hesitant. I had only been to Japan once or twice before, and had been told on numerous occasions that Japanese students had no initiative, did not not want to make choices, and were not interested in learning-how-to-learn. I asked the people who brought me to Japan, "Are you sure you want me to talk about learner-centeredness?" They assured me that they did. In the event, I was delighted with the experience. Here are some of the ways in which teachers in Japan involved learners in the learning process. I have recorded the verbatim statements of the seminar participants who were Japanese and American teachers of English in Japan working with secondary students and adults:

I tell the students what they are going to learn and why.

I now tell my students how they are going to learn and why.

My students love to make choices and come to decisions - for example, if they want to read or listen first.

I get my students to monitor and evaluate their language use.

I teach my students to monitor and evaluate their progress.

My students evaluate ways of learning in class which seem to work best for them.

I get my students to think about ways of using English outside of the classroom.

My students predict how much they will have learned by the end of the course.

At the beginning of the semester, I ask my students to predict what they will be able to do at the end of the semester.

My students get lots of practice in identifying their own errors.

I get my advanced level students to find alternative ways of doing a task.

I make my students responsible for various aspects of classroom management.

This year, I have been experimenting with alternative ways of grouping learners.

I've been training my students to set their own goals and objectives.

My students have been involved in setting up and running a self-access centre.

Many of the materials in my classroom have been adapted or created by the learners.

I try and make my learners aware of the learning strategy implications of classroom tasks.

I help my learners see similarities and differences between learning language and learning other subjects.

My students have become very good at evaluating the difficulty of objectives and classroom tasks.

These studies illustrate the point made earlier, that learner/learning-centeredness are not all-or-nothing concepts, and that there are many ways in which the philosophy can be activated, even in classrooms which are directed by externally mandated curricula. At a modest level, simply telling learners what they are going to learn, how they are going to learn, and why they are going to engage in particular activities will make a classroom more learner-centered. All other things being equal, classrooms where teachers spell out their objectives will be more learner-centered than those where teachers do not explain program objectives. At the other end of the scale are classes where learners are trained to set their own objectives. This represents a very advanced application of the concept.

I am currently involved in developing an international textbook series for an American publisher in which learners are systematically trained in applying learning, communication, and cognitive strategies to the process of learning language. This focus helps to realize the

principle set out in *The Learner-Centered Curriculum* in which I argued that

> A learner-centered approach will be concerned with the development of and sensitizing of learners to their role as learners. In consequence, while one set of goals and objectives will relate to the development of language skills, a parallel set will relate to the development of learning skills. (Nunan, 1988, p. 179).

Examples of these tasks are set out in *Table 3*. As indicated above, these are designed to develop both language ability and learning skills so that the learners can go on learning independently beyond the language classroom. Although this project is at an early stage, field test feedback has been extremely encouraging. For additional ideas on ways in which these strategies can be exploited in the language, classroom, see Nunan and Lamb (forthcoming).

Table 3

STRATEGY	GLOSS / EXAMPLE
Classifying	Putting things that are similar together in groups. *Example: Study a list of names and segregate them into male and female.*
Predicting	Predicting what is to come in the learning process. *Example: Look at unit title and objectives and infer what will be learned.*
Brainstorming	Thinking of as many new words and ideas as one can. *Example: Work in a group and think of as many occupations as you can.*
Selective Listening	Listening for key words and information without trying to understand every word. *Example: Listen to a conversation and identify the number of speakers.*

(Table cont'd)

Table 3 *(cont'd)*

Personalizing	Learners share their own opinions, feelings and ideas about a subject. *Example: Read a letter from a person in trouble and say what you would do.*
Scanning	Looking quickly through a written text for specific information. *Example: Look at the text and decide if it is a newspaper article, a letter, or an advertisement.*
Conversational Patterns	Using expressions to start conversations and keep them going. *Example: Match formulaic expressions to situations.*
Cooperating	Sharing ideas and learning with other students. *Example: Work in small groups to read a text and complete a table.*
Discovering	Looking for patterns in language. *Example: Study a conversation and note the rule for forming the simple past tense.*
Practicing	Doing controlled exercises to improve knowledge and skills. *Exercise: Listen to a conversation, and conduct it with a partner.*
Role playing	Pretending to be somebody else and using the language for the situation you are in. *Example: You are a reporter. Use the information from the reading to interview the writer.*

(From Nunan, 1992c)

Conclusion

In this article, I have revisited the notions of learner- and learning-centeredness. I have argued that a learner-centered curriculum will contain similar elements to those contained in traditional curriculum development; that is, needs analysis, goal and objective setting, methodology, materials development, assessment and evaluation. However, the key difference between learner-centered and traditional curriculum development is that in the former the curriculum is a collaborative effort between teachers and learners, because learners will be more closely involved in the decision-making process regarding the content of the curriculum, how it is taught, and how it is assessed. I further argued that the concept is not an absolute one, and that it does not assume that learners will be able to make choices from the beginning of instruction. In fact, parallel with the learning of language, learners should also be taught about the learning process. Where feasible, while one major aim or set of aims will relate to the teaching of specific language skills, other aims will relate to the development of learning skills. Such aims may include the following:

- To provide learners with efficient learning strategies.
- To assist learners to identify their own preferred ways of learning.
- To develop skills needed to negotiate the curriculum.
- To encourage learners to set their own objectives.
- To encourage learners to adopt realistic goals and time frames.
- To develop learners' skills in self-evaluation. (Nunan, 1988, p. 3)

In the second part of the paper, I reviewed some recent empirical research which has implications for curriculum development. From the research reviewed in this paper, a number of conclusions emerge, specifically that:

- Learners are different and learn in different ways.
- There may be as many different interpretations of what is going on in the classroom as there are learners in the class.
- Learning is enhanced when learners are actively involved in communicating with each other and with the teacher in the target language.
- Learning is enhanced when learners are given opportunities to select content and learning tasks and also when they are provided with opportunities to evaluate their own progress.

Following this review, I provided some practical suggestions which have been developed and tried in Japan. The paper concluded with a brief look at a new curriculum which is organized around a set of cognitive, communicative learning strategies.

References

Assinder, W. (1991). Peer teaching, peer learning: one model. *ELT Journal, 45, 3, 218-229.*

Brindley, G. (1989). *Assessing achievement in a learner-centered curriculum.* Sydney: NCELTR.

Cathcart, R., & Vaughn, S. *Real conversations: Beginning listening and speaking activities.* Boston: Heinle & Heinle.

Dam, L., & Gabrielsen, G. (1988). Developing learner autonomy in a school context. A six-year experiment beginning in the learner's first year of English. In H. Holec (Ed.), *Autonomy and self-directed learning: Present fields of application.* Strasbourg: Council of Europe.

Ellis, G. & Sinclair, B. (1989). *Learning to learn English: A course in learner training.* Cambridge,: Cambridge University Press.

Hall, S. (1991). The effect of split information tasks on the acquisition of mathematical vocabulary. Unpublished dissertation, Victoria University of Wellington, Wellington, New Zealand. In D. Nunan (Ed.), (1992a). *Collaborative language learning and teaching.* Cambridge: Cambridge University Press.

Kohonen, V. (1992). Experiential language learning: Second language learning as cooperative learner education. In D. Nunan (Ed.), (1992a). *Collaborative language learning and teaching* Cambridge: Cambridge University Press.

Kumaravadivelu, B. (1991). Language-learning tasks: Teacher intention and learner interpretation. *ELT Journal, 45, 2, 98-107.*

Legutke, M. and Thomas, H. (1991). *Process and experience in the language classroom.* London: Longman.

Lim, S. (1992). Investigating learner participation in teacher-led classroom discussions in junior colleges in Singapore from a second language acquisition perspective. Unpublished doctoral dissertation, National University of Singapore.

Nunan, D. (1988). *The learner-centered curriculum.* Cambridge: Cambridge University Press.

Nunan, D. 1991. *Language teaching methodology.* London: Prentice Hall.

Nunan, D. (Ed.). (1992a). *Collaborative language learning and teaching.* Cambridge: Cambridge University Press.

BEST COPY AVAILABLE

Nunan, D. (1992b). *Research methods in language learning*. Cambridge: Cambridge University Press.

Nunan, D. (1992c). *Atlas. A course in learning-centered communication*. Level II. *Pilot Edition*. Boston: Heinle & Heinle.

Nunan, D. & Lamb, C. (forthcoming) *Communication tasks in language learning*.

Scarcella, R. & Oxford, R. (1992). *The tapestry of language learning: The individual in the communicative classroom*. Boston: Heinle & Heinle.

Slimani, Y. (1991). Evaluating classroom interaction. In J. C. Alderson and A. Beretta (Eds.). *Evaluating second language education*. Cambridge: Cambridge University Press.

Widdows, S. & P. Voller. (1991). PANSI: A survey of ELT needs of Japanese university students. *Cross Currents*, XVIII, 2.

Willing, K. (1988). *Learning styles in adult migrant education*. Adelaide: National Curriculum Resource Centre.

Author

DAVID NUNAN, National Centre for English Language Teaching and Research, Macquarie University, Sydney, New South Wales 2109, Australia. Specializations: learner-centered curriculum, language teaching methodology.

Applied Language Learning
1993, Vol. 4, Nos. 1 & 2, pp. 19-54

Proficiency and Performance in Language Testing

James Child
Department of Defense

Ray T. Clifford
Defense Language Institute

Pardee Lowe, Jr.
CIA Language School

The purpose of this article is to establish working definitions of proficiency and performance in language testing. Following the definition of these terms, their role in multidimensional testing is examined. Then, the relationship between proficiency and performance is reviewed. Next, similarities and differences between the two types of tests are identified and discussed. Also, the two test types are examined vis-à-vis the proficiency trisection. Finally, performance test scoring dilemmas and topics requiring further research are identified.

Introduction

The continuing lack of agreement in the language community at large on what constitutes "language proficiency" has led to the predictable impasse that accompanies the use of any term associated with a range of interpretations. Larson and Jones (1984) make this point when they cite definitions of proficiency ranging from a vaguely defined ability to use the language, through the specific and detailed skill level descriptions developed by the Interagency Language Roundtable (ILR), the Educational Testing Service (ETS), and the American Council on the Teaching of Foreign Languages (ACTFL).[1] These descriptions originated at the US State Department's Foreign Service Institute (FSI) in the mid-1950s and have since been adapted government-wide. The ILR Testing Committee has traditionally been charged with their revision and further specification.[2] In the early 1980s, the ETS held a conference on the Common Metric. In 1982, ACTFL, in conjunction with ETS, proposed a set of Provisional Guidelines derived from and commensurate with the ILR Skill Level Descriptions. These were revised by ACTFL in 1986. (See *Appendices 1* and 2 for the ACTFL guidelines and the ILR descriptions.)

Scholars and scientists have always been faced with performance demands in drawing on second language skills in their research in any of a number of fields, the languages for the most part being French, German, English, and, earlier, Latin. More recently there has been a tremendous increase in the importance of Russian on the world stage, followed by languages such as Japanese and Arabic. In today's work world, global economic and security standards are the driving forces for performance in the four language skills. Thus, the need for testing functional foreign language based on the ACTFL / ETS / ILR (AEI) definitions has been widely recognized.

Concepts of Proficiency and Performance

In this article, *proficiency* is considered as applying to an unspecified or a potential product, and a second term, *performance* (or more accurately, "mission performance"), is introduced to apply to a specific, defined product.[3] It may sometimes be impractical to draw an absolute line between these categories. Nevertheless, the dichotomy may be a useful one in helping applied linguists in making required judgments and managers in making administrative decisions.

The term *performance* relates to a person's actual skill level in performing a language task typical of or similar to one required in the workplace or in any other milieu where a "product" of some kind is called for. Performance is not to be confused with the same term in Chomsky's famous competence/performance distinction (Chomsky, 1975).

The notion of proficiency is difficult to define. Depending on one's definition, the term can be used: (a) in the more general sense, for any textual application of authentic materials, or (b) in the more restricted sense, for highly specialized content areas. In the latter case, proficiency simply subsumes the concept of performance. Any formal applications focusing exclusively on the ability to use proper grammatical forms in isolation (e.g., to change nouns from singular to plural in a sentence) are beyond the notion of proficiency. Thus, this notion is generally associated with the skill of the learner in applying the formal structures he or she controls to convey messages of a non-specialized nature processed in the production or reception modalities. An individual's skill levels are ordinarily inferred from test results that are based on this model. In government-sponsored proficiency testing, few or no aids are made available, and time limits are imposed. These constraints are partially offset by the absence of a requirement for highly specialized or arcane knowledge on the part of the testee, who is neither at an advantage nor a disadvantage vis-à-vis difficult subject matter. It is the former, or government, view of general proficiency that is applied here.

Proficiency (measured by formal testing or informally inferred from observation) is person-directed: Individual X is—based on the evidence available—able to process (in production or reception mode, as applicable) a variety of texts at level Y, depending on the person's range of general and technical knowledge. The individual's performance, on

the other hand, is judged according to the degree of his or her success in carrying out a formal task involving the language in question, with the understanding that the task in question is, or easily could be, fulfilling a required job, and is not a sampling of language behavior.

Thus, performance is carried out one text at a time, and is very often limited to one or a few content areas associated with some kind of job. For example, control of the subject matter (generally acquired in one's own language) in one of a few content areas frequently carries the day, allowing the speaker or reader to feed his or her knowledge of the subject matter at hand back into the second language text in the attempt to communicate or to understand the text. Successful performance is thus no guarantee of the user's control of the second language, for a switch to another content area (whether in the production or the reception modalities) may well lead to a disastrous performance. It is true that performance itself need not be couched in absolute terms: A work situation may dictate that a subpar performance is better than nothing.

Absolute standards of both proficiency and performance are, however, necessary as guidelines against which language tests can be designed and examinees can be evaluated to give testers, teachers, managers, and others concerned a solid metric on which to base ratings. Both types of tests are used by the government.

Relationships between Proficiency and Performance

There are situations in which proficiency and performance overlap, or proficiency encapsulates performance. (See *Figure 1*.) Even in such cases, though, fuller examination of the arenas in which the terms are relevant helps bring the distinction into focus. Thus, a reasonably high level of proficiency means that the language user is comfortable reading about or discussing a wide number of topics of a general nature and a smaller number of more specialized topics familiar to him or her. And, if the user is at a very high level, he or she can deal to some degree with specialized topics of an unfamiliar nature, even though there may be text segments imperfectly understood in the reception modalities or expressed quaintly or inexactly in the production modalities.

Figures 1-3 illustrate the most common relationships between proficiency and performance testing domains. Although it is theoretically possible that a performance test domain might fall entirely outside the general domain of proficiency testing (*Figure 2*), this situation would normally only occur when the proficiency domain is restricted to Level 1 functions and contexts. At higher levels of proficiency, it is more likely that a performance test would either be subsumed under the broad proficiency domain (*Figure 1*), or overlap with, and extend beyond, the proficiency domain into a specialized area of competence (*Figure 3*).

Testing Domains
Figure 1

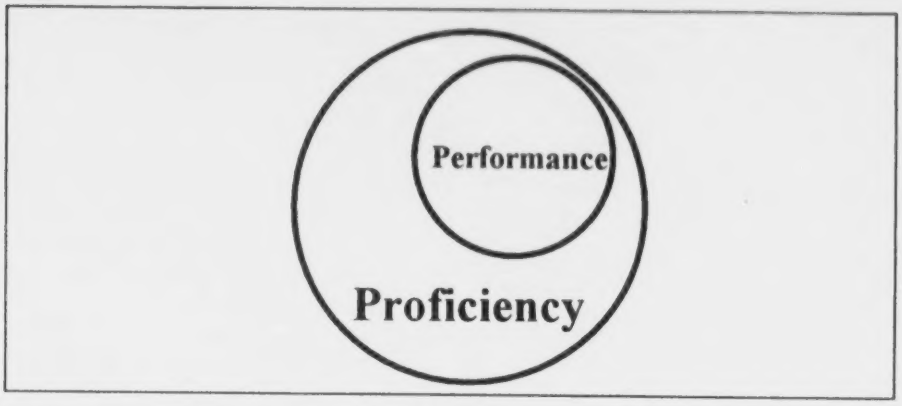


Figure 2

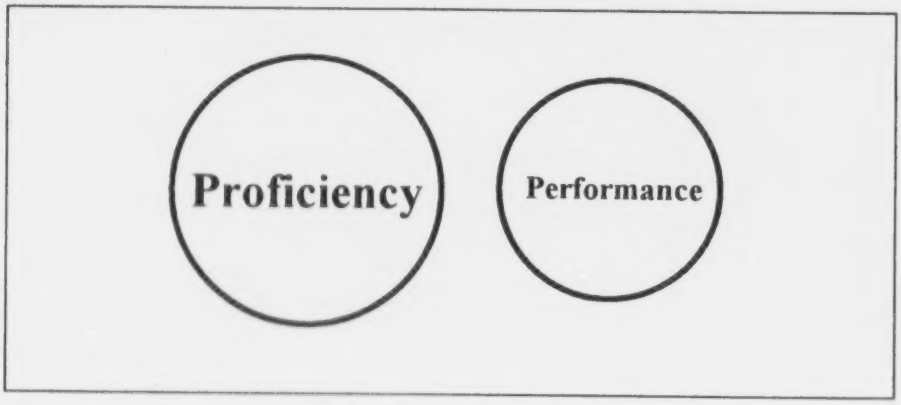
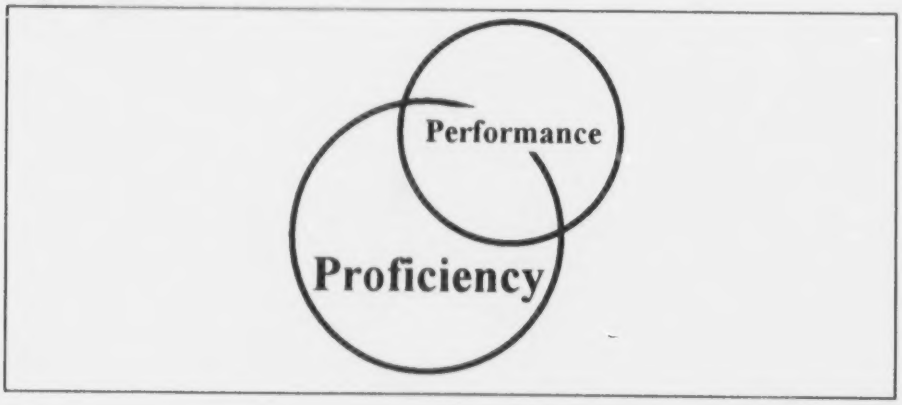


Figure 3



Similarities and Differences between Proficiency and Performance Tests

Although there are many similarities between proficiency and performance tests, there are also significant differences that lead to divergent results and different interpretations of scores. The main similarity, of course, is that both types of tests prompt or elicit responses within a communicative framework. No test or subtest of either test type would, for example, require the testee to list variant forms of the genitive case or even indicate deep-to-surface valency or case relationships in a set of sentences. In other words, unlike achievement testing, both a proficiency test and a performance test get at a candidate's ability to *use* the language rather than at their knowledge of facts *about* a linguistic system. The heart of the matter is captured under the rubric of language use, not language description.

The differences between the two test types can be best defined in terms of individual capabilities in processing texts of a general nature at various levels versus the requirements imposed by a job (usually professional, but sometimes personal) involving specific text types falling into a relatively constricted range (say, Levels 2 through low Level 4). In assessing a subject's proficiency, the intent is to determine the highest point on an ascending scale at which he or she can function. To do this, the tester provides a sufficient variety and number of tasks—as Child (1987) suggests: from Level 1, Orientational, to Level 4, Projective (or Level 5, Special Purpose).

Performance testing, on the other hand, is oriented to a current or prospective job. Candidates taking government performance tests of listening comprehension may well be required to demonstrate competence at more than one level (generally between 2+ and 3+), but the critical concentration is on comprehension of the designated subject matter. In proving their mastery, the candidates are not only permitted but encouraged to use reference materials of various kinds during test sessions. Since they presumably have at least some background in the content area tested, they can be expected to use such knowledge as part of their overall comprehension strategy.

In testing the individual's proficiency level in a form/content paradigm, language forms are tested as vehicles to convey messages rather than as formal structures to be learned for their own sake. (It should be noted that the seeming priority accorded to mastery of form is, therefore, a question of perspective in de- and re-contextualizing, and not a matter of a typologically- or historically-oriented paradigm). For example, testees may have to move quickly or shift gears in processing texts, but the texts in these proficiency tests themselves make no heavy subject matter demands.

Performance measures, on the other hand, are central to the content/form paradigm (Child, 1990).[4] Performance measures assume that the testee has the specialized knowledge to carry out an actual job assignment. Far from being ruled out or neutralized, subject matter control is an essential skill which the testee must demonstrate, using the best working aids available to support his or her mission. It is, of course, true, as was noted earlier, that there is no absolute line of demarcation between proficiency and performance: An account of a current event such as a border clash or political coup may be used either to tap general functions and content in a second language ("narration of factual content," Level 2, p. 50, *ILR Level Descriptions*) or to elicit specific background knowledge needed for a briefing or report.

Candidates doing performance tests are not usually expected to be able to cope with just any unspecified topic, cultural or technical: They often have some notion in advance of the kinds and amount of material they are to handle. This information enhances the semantic feedback process. Thus, semantic feedback may operate as only an additive factor for some test takers. *Figure 4* offers a visual image of the performance paradigm: a pyramid, the base of which is Level 0 and the apex Level 5, with subject matter knowledge superimposed on, but not totally confined within, the form itself.

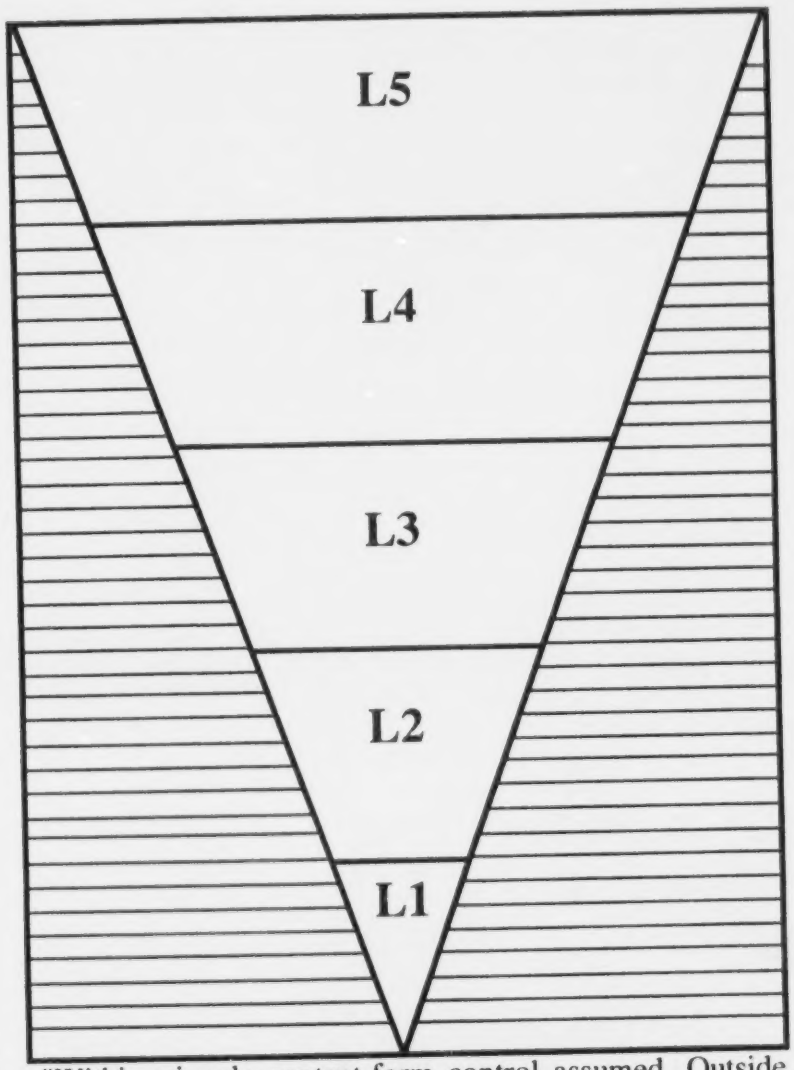
Features of Proficiency and Performance Testing

The key features of proficiency and performance testing are highlighted when these testing models are viewed from the perspective of the three major components of the proficiency trisection: accuracy, context, and function (Higgs & Clifford, 1982).[5] See *Figure 5*. The validity of this trisectonal model of proficiency has been confirmed by data collected by de Jong and van Ginkel (1992, pp. 202-203). Each of these components will be considered in turn:

ACCURACY. The accuracy component can be and usually is identical for both testing models, so this component poses no significant challenge when moving from one testing system to the other.

CONTEXT/TOPICS. For the reasons outlined above, proficiency and performance testing models diverge most in the subject matter addressed. In general, a proficiency test covers a wide range of general topics, and the score obtained serves as an indicator of general ability. Managers can extrapolate or infer from this indicator the ability to handle tasks across a range of potential jobs. However, because the test candidate may not yet have mastered the lexicon or developed a cognitive understanding of a particular job, a proficiency score of Level 2 should be interpreted as a prerequisite, but not as a guarantee of Level 2 working ability for a specific job.

Figure 4
Performance Paradigm



"Within triangle content-form control assumed. Outside traingle content knowledge an added requirement which, once attained, will enable content-form control."

Figure 5
Functional Trisection of Oral Proficiency Levels

LEVEL	FUNCTIONS	CONTENT*	ACCURACY**
FSI Speaking Level	Task accomplished, Attitudes expressed, Tone conveyed.	Topics, subject areas, activities and jobs addressed.	Acceptability, quality, and accuracy of message conveyed.
5	Functions equivalent to an Educated Native Speaker (ENS).	All subjects.	Performance equivalent to an ENS.
4	Able to tailor language to fit audience, counsel, persuade, negotiate, represent a point of view, and interpret for dignitaries.	All topics normally pertinent to professional needs.	Nearly equivalent to an ENS. Speech is extensive, precise, appropriate to every occasion with only occasional errors.
3	Can converse in formal and informal situations, resolve problem situations, deal with unfamiliar topics, provide explanations, describe in opinions, and hypothesize.	Practical, social, professional, and abstract topics, particular interests, and special fields of competence.	Errors never interfere with understanding and rarely disturb the native speaker (NS). Only sporadic errors in basic structures.
2	Able to fully participate in casual conversations, can express facts, give instructions, describe, report on, and provide narration about current, past, and future activities.	Concrete topics such as own background, family, and interests, work, travel, and current events.	Understandable to an NS *not* used to dealing with foreigners; sometimes miscommunicates.
1	Can create with the language: ask and answer questions, participate in short conversations.	Everyday survival topics and courtesy requirements.	Intelligible to an NS used to dealing with foreigners.
0	No functional ability.	None.	Unintelligible.

*May be job specific.
**See also factor performance rating scales.

A performance test covers a selected, restricted domain and the results can be interpreted as an indication of one's ability to perform tasks similar to those selected for the test sample. However, because of the restricted focus of the test items, the results cannot be generalized to imply commensurate ability across a wide range of topics or jobs.

FUNCTION. The application of the functional category of the trisection to performance testing is in principle straightforward but in practice complex. Theoretically, the same hierarchical progression of task complexity found in the proficiency scale can be applied to performance testing, and in test design it usually is.

As the topics become known or even predictable, a confounding of performance variables occurs which hinders explanatory interpretation of test results. Thus, due to the restricted domain of the test, the performance observed may be greatly influenced by familiarity with the subject, or in Child's terms, semantic feedback (1987, p. 105). To avoid this problem, the joint Matriculation Board examiners in the U.K. have determined that test content should be on topics that nobody taking the test should be familiar with, such as bookbinding and the raising of ancient sunken ships (Douglas, personal communication). Furthermore, if tests used for creating test items are selected on the basis of frequency of occurrence in the job rather than criticality to the job, then the scores obtained are misleading.

Proficiency and Performance Tests in Government Use

Government language tests are generally neither purely proficiency nor performance measures but some mixture of the two, with one or the other approach predominating. It may, therefore, be useful to discuss some tests widely used in government agencies and tentatively assign them to one of the two categories, proficiency or performance.

The Oral Proficiency Interview (OPI) enjoys great favor in government testing circles, and with good reason. Although highly labor intensive, it can usually be counted on to give teaching departments or managers in the language workplace a highly accurate profile of the language learner or user with respect to skill level attainment. Ranging as it does over a spectrum of Speaking-0 to Speaking-5, it cannot be expected to address particular subject matter areas in any detail, but it does require of the testee flexibility and breadth as it moves from one topic to another. OPI testers avoid to the extent possible specialized cultural issues or highly technical discourses so that testees will neither be at a disadvantage dealing with unfamiliar content nor will they have any advantage by using content knowledge as a mask for second language weaknesses. For these reasons, the OPI is the quintessential proficiency test.

Much the same can be said of the *Computer-Adaptive Tests* (CATs) of reading proficiency now rapidly gaining in popularity, as well as pre-computer models of standardized, objectively scored tests used at the Defense Language Institute (DLI) for both reading and listening, the *Defense Language Proficiency* Tests (DLPT). Unlike the OPI, these test batteries address the reception modalities; like the OPI, they seek to neutralize the demands of highly specialized content. Naturally, the test items are meant to elicit the examinee's understanding of the content of various passages at different levels, but they are so structured that he or she does not have to bring extensive technical knowledge to the task. The same can be said of other kinds of tests in use throughout government, often "quasi-translation" reading tests where the elicitation of general skill rather than particular knowledge is the aim of the exercises. The OPIs, CATs, DLPTs, and "quasi-translation" tests meet the definition of "proficiency" tests.

Two examples of performance models are: one of the two parts of a Department of Defense Level 3 examination that includes a job-oriented exercise requiring specialized knowledge and skills; and certain Foreign Service Institute (FSI) tests more or less geared to prospective field assignments for employees. Performance examinations—both spoken and written—may span a considerable range, from "quasi-performance" tests based on texts similar but not identical to those associated with jobs, to the actual production of job samples (Clark, 1979).

Scoring Dilemmas

Even with well-designed and constructed performance tests, the interpretation of scores can become problematic. The central issue that must always be formally addressed in performance testing is the contrast between *frequency* and *criticality* of texts used for creating test items. When the text is selected based on frequency of occurrence, then the test has high face validity—but a high score will not indicate an ability to successfully handle less frequent but critical job tasks. When texts are chosen on the basis of criticality, then a high score is indicative of an individual's ability to complete critical tasks successfully—but the test will be criticized as non-representative of the real day-to-day job.

The practical solution to this dilemma is to address frequency and criticality separately by including at least two independently scored subtests within each performance test. This approach satisfies face validity concerns and the independent scores from these subtests provide managers with a much clearer picture of each test candidate's job performance abilities.

Familiarity with the job will always serve as an advance organizer for the test situation and it may significantly increase scores on purported language tests, just as a "hothouse special," a pet topic or

two of a candidate, may demonstrate increased performance on rehearsed topics within a proficiency test (DLI, 1982, pp. 4-19). Thus, these candidates' scores on a proficiency test may be inflated as they respond on the basis of prior knowledge or experience rather than on information gleaned from the language reading or listening text at hand. For example, since most job communications follow predictable exchange patterns, knowing the situation allows one to predict the majority of these formulaic communicative exchanges. "How are you?" is routinely followed by "Fine, thanks," and seldom by "Get lost!" Similarly, adult test candidates would likely be able to answer a question about the freezing point of water, whether or not they had understood a passage on that topic in a foreign language chemistry textbook. However, a distinction between superficial and deeper knowledge may be made (Douglas, 1993). For example, an article on the ozone layer in *Time* differs from articles on the same topic in *Scientific American* or in the *Journal of Atmospheric Sciences*.

There are probably two different types of language which may appear here to confound the performance/proficiency distinction and thus contribute to the scoring dilemmas, namely, *esoteric* and *extrinsic* use of language. *Esoteric* is used here in the sense of "some skill, knowledge, etc., limited to a small group." An example of this type of language use would be reading voice prints, something people generally proficient in a language would never be asked to do and would probably always have to be specially taught and which should be tested by a performance test.

Moreover, there are tasks that are *extrinsic* to a lower level of proficiency, but would be intrinsic, that is, doable by someone at a higher level. Such a task might be reading a telegram in German where the grammar differs from that customary at the lower levels of the language. Naturally, higher level readers could accomplish such a task, so it is not clear how this second category should be tested—by performance tests for lower level readers or by higher level proficiency tests.

These two categories, areas esoteric to general proficiency at any level and areas extrinsic to some lower level of proficiency, but intrinsic to a higher one, as well as the requirements they place on testing the performance/proficiency distinction, deserve more elaboration than space here permits if a complete testing model is to be developed.

Conclusion

In this article, a working distinction between performance and proficiency has been posited in general terms. The major difference between proficiency testing and performance testing is generally in the selection of the texts domain. This difference in test domain can result

in significant discrepancies between performance and proficiency test results, and these differences in test scores generally relate to differences in the application of these two testing models.

Thus, if one wants an indication of an individual's general language ability—either for application across a wide variety of jobs or in situations where the test candidates have not yet been taught their job skills—then a proficiency test is the appropriate instrument. On the other hand, if one desires an indication of an individual's ability to perform a specific job, then one needs a job performance test. However, the narrower, specialized focus of performance testing also creates special requirements for test design and score reporting. For instance, for accurate decision-making, performance tests must provide independent scores for the test candidates' ability to handle high-frequency, routine tasks and for their ability to handle critical, albeit infrequent, job tasks. It is hoped that this delineation between proficiency and performance testing may both inform current testing practices and motivate continued research into this important distinction.

Notes

1. For a more in-depth discussion of the ACTFL, ETS, and ILR (AEI) scales and their special characteristics, see Lowe (1985 and 1986). The former article discusses the early research carried out on the scales; the latter delves into critics' views of problems with defining and applying the scales.

2. ILR is an informal organization on which may sit all government agencies interested in foreign language training, testing, and employing federal employees who need foreign language on the job. Traditionally, the ILR has had committees concerned with managing federal foreign language resources, with training and curriculum design, with doing research, with implementing technology for training and testing, and with assessing the foreign language proficiency of prospective and current federal employees.

3. In general, the naming can proceed in one of two ways. The more common approach simultaneously names and defines a concept, usually as narrowly as possible, and then works outward, seeking to identify the named concept's referents in the real world. In contrast, the approach most often used in ILR work begins with general observation. In the AEI definitions, some reality often long worked with—is named. Then,

one works inward from the named reality to ascertain all that the concept encompasses. This second approach frequently means that the fuller definition considerably postdates the act of naming. In fact, this process often lasts years, as the ongoing refining of the ILR Speaking definitions for Oral Proficiency Interview (OPI) testing illustrates. In this article, the second approach has been followed, because it outlines large significant areas with which language testers must deal. A fuller elaboration is necessary, but must await later research and explication.

4. The form/content versus content/form distinction (Child, 1990) focuses in each case on the primacy of the former member of the pair over the latter. Thus, those learning a language are often led to stress the meaning of the form rather than the meaning of the overall content. As a result, they miss meaning that extends beyond the word or the phrase. On the other hand, those approaching target language meaning from the content/form perspective often describe meaning from the overall context (nature of the topic, semantic feedback, logic, etc.), either sensing or knowing that some statement must be true, because it matches the content required. Where both the subject matter and the language are brand new, one tends to be thrown back onto a form/content approach.

5. The trisection referred to here is the one for oral proficiency created by Ray T. Clifford and Pardee Lowe, Jr., reproduced in Higgs and Clifford (1982). Available within the ILR are similar trisections for Listening by Carl Ericson of DLI and for Reading by Ray T. Clifford, Pardee Lowe, Jr., and Martha Herzog (Lowe, 1984, pp. 387-388). Due to the different nature of a reception skill, there also exists a Reading Quintasection.

6. The *ILR Handbook on Oral Interview Testing* provides a definition for "Hot-House Hank," the candidate who has a pet topic or two upon which he expatiates with élan and éclat, but who fails to use higher level vocabulary and grammar on other topics posed by the interviewers (p. 4-19).

Appendix 1
ACTFL Proficiency Guidelines *

The 1986 proficiency guidelines represent a hierarchy of global characterizations of integrated performance in speaking, listening, reading, and writing. Each description is a representative, not an exhaustive, sample of a particular range of ability, and each level subsumes all previous levels, moving from simple to complex in an "all-before-and-more" fashion.

Because these guidelines identify stages of proficiency, as opposed to achievement, they are not intended to measure what an individual has achieved through specific classroom instruction but rather to allow assessment of what an individual can and cannot do, regardless of where, when, or how the language has been learned or acquired; thus, the words *learned* and *acquired* are used in the broadest sense. These guidelines are not based on a particular linguistic theory or pedagogical method, since the guidelines are proficiency-based, as opposed to achievement-based, and are intended to be used for global assessment.

The 1986 guidelines should not be considered the definitive version, since the construction and utilization of language proficiency guidelines is a dynamic, interactive process. The academic sector, like the government sector, will continue to refine and update the criteria periodically to reflect the needs of the users and the advances of the profession. In this vein, ACTFL owes a continuing debt to the creators of the 1982 provisional proficiency guidelines and, of course, to the members of the Interagency Language Roundtable Testing Committee, the creators of the government's Language Skill Level Descriptions.

ACTFL would like to thank the following individuals for their contributions on this current guidelines project:

Heidi Byrnes
James Child
Nina Levinson
Pardee Lowe, Jr.
Seiichi Makino
Irene Thompson
A. Ronald Walton

These proficiency guidelines are the product of grants from the U.S. Department of Education.

Generic Descriptions–Speaking

Novice	The Novice level is characterized by an ability to communicate minimally with learned material.
Novice-Low	Oral production consists of isolated words and perhaps a few high-frequency phrases. Essentially no functional communicative ability.

* Canale, M. & Byrnes, H. (Eds.). (1987). Defining and developing proficiency: Guidelines, implementations, and concepts (pp. 15-24). Lincolnwood, IL: National Textbook. Reprinted by permission of the Publisher and the American Council on the Teaching of Foreign Languages.

BEST COPY AVAILABLE

Novice-Mid	Oral production continues to consist of isolated words and learned phrases within very predictable areas of need, although quantity is increased. Vocabulary is sufficient only for handling simple, elementary needs and expressing basic courtesies. Utterances rarely consist of more than two or three words and show frequent long pauses and repetition of interlocutor's words. Speaker may have some difficulty producing even the simplest utterances. Some Novice-Mid speakers will be understood only with great difficulty.
Novice-High	Able to satisfy partially the requirements of basic communicative exchanges by relying heavily on learned utterances but occasionally expanding these through simple recombinations of their elements. Can ask questions or make statements involving learned material. Shows signs of spontaneity, although this falls short of real autonomy of expression. Speech continues to consist of learned utterances rather than of personalized, situationally adapted ones. Vocabulary centers on areas such as basic objects, places, and most common kinship terms. Pronunciation may still be strongly influenced by first language. Errors are frequent and, in spite of repetition, some Novice-High speakers will have difficulty being understood even by sympathetic interlocutors.
Intermediate	The intermediate level is characterized by an ability to —create with the language by combining and recombining learned elements, though primarily in a reactive mode; —initiate, minimally sustain, and close in a simple way basic communicative tasks; and —ask and answer questions.
Intermediate-Low	Able to handle successfully a limited number of interactive, task-oriented and social situations. Can ask and answer questions, initiate and respond to simple statements, and maintain face-to-face conversation, although in a highly restricted manner and with much linguistic inaccuracy. Within these limitations, can perform such tasks as introducing self, ordering a meal, asking directions, and making purchases. Vocabulary is adequate to express only the most elementary needs. Strong interference from native language may occur. Misunderstandings frequently arise, but with repetition, the Intermediate-Low speaker can generally be understood by sympathetic interlocutors.
Intermediate-Mid	Able to handle successfully a variety of uncomplicated, basic and communicative tasks and social situations. Can talk simply about self and family members. Can ask and answer questions and participate in simple conversations on topics beyond the most immediate needs; e.g., personal history and leisure-time activities. Utterance length increases slightly, but speech may continue to be characterized by frequent long pauses, since the smooth incorporation of even basic conversational strategies is often hindered as the speaker struggles to create appropriate language forms. Pronunciation may continue to be strongly influenced by first language and fluency may still be strained. Although misunderstandings still arise, the Intermediate-Mid speaker can generally be understood by sympathetic interlocutors.
Intermediate-High	Able to handle successfully most uncomplicated communicative tasks and social situations. Can initiate, sustain, and close a general conver-

sation with a number of strategies appropriate to a range of circumstances and topics, but errors are evident. Limited vocabulary still necessitates hesitation and may bring about slightly unexpected circumlocution. There is emerging evidence of connected discourse, particularly for simple narration and/or description. The Intermediate-High speaker can generally be understood even by interlocutors not accustomed to dealing with speaking at this level, but repetition may still be required.

Advanced

The Advanced level is characterized by an ability to
—converse in a clearly participatory fashion;
—initiate, sustain, and bring to closure a wide variety of communicative tasks, including those that require an increased ability to convey meaning with diverse language strategies due to a complication or an unforeseen turn of events;
—satisfy the requirements of school and work situations; and
—narrate and describe with paragraph-length connected discourse.

Advanced

Able to satisfy the requirements of everyday situations and routine school and work requirements. Can handle with confidence but not with facility complicated tasks and social situations, such as elaborating, complaining, and apologizing. Can narrate and describe with some details, linking sentences together smoothly. Can communicate facts and talk casually about topics of current public and personal interest, using general vocabulary. Shortcomings can often be smoothed over by communicative strategies, such as pause fillers, stalling devices, and different rates of speech. Circumlocution which arises from vocabulary or syntactic limitations very often is quite successful, though some groping for words may still be evident. The Advanced-level speaker can be understood without difficulty by native interlocutors.

Advanced-Plus

Able to satisfy the requirements of a broad variety of everyday, school, and work situations. Can discuss concrete topics relating to particular interests and special fields of competence. There is emerging evidence of ability to support opinions, explain in detail, and hypothesize. The Advanced-Plus speaker often shows a well-developed ability to compensate for an imperfect grasp of some forms with confident use of communicative strategies, such as paraphrasing and circumlocution. Differentiated vocabulary and intonation are effectively used to communicate fine shades of meaning. The Advanced-Plus speaker often shows remarkable fluency and ease of speech, but under the demands of Superior-level, complex tasks, language may break down or prove inadequate.

Superior

The Superior level is characterized by an ability to
—participate effectively in most formal and informal conversations on practical, social, professional, and abstract topics; and
—support opinions and hypothesize using native-like discourse strategies.

Superior

Able to speak the language with sufficient accuracy to participate effectively in most formal and informal conversations on practical, social, professional, and abstract topics. Can discuss special fields of competence and interest with ease. Can support opinions and hypothesize, but may not be able to tailor language to audience or discuss in

depth highly abstract or unfamiliar topics. Usually the Superior-level speaker is only partially familiar with regional or other dialectical variants. The Superior-level speaker commands a wide variety of interactive strategies and shows good awareness of discourse strategies. The latter involves the ability to distinguish main ideas from supporting information through syntactic, lexical, and suprasegmental features (pitch, stress, intonation). Sporadic errors may occur, particularly in low-frequency structures and some complex high-frequency structures more common to formal writing, but no patterns of error are evident. Errors do not disturb the native speaker or interfere with communication.

Generic Descriptions–Listening

These guidelines assume that all listening tasks take place in an authentic environment at a normal rate of speech using standard or near-standard norms.

Novice	The Novice level is characterized by an ability to recognize learned material and isolated words and phrases when strongly supported by context.
Novice-Low	Understanding is limited to occasional isolated words, such as cognates, borrowed words, and high-frequency social conventions. Essentially no ability to comprehend even short utterances.
Novice-Mid	Able to understand some short, learned utterances, particularly where context strongly supports understanding and speech is clearly audible. Comprehends some words and phrases from simple questions, statements, high-frequency commands, and courtesy formulae about topics that refer to basic personal information or the immediate physical setting. The listener requires long pauses for assimilation and periodically requests repetition and/or a slower rate of speech.
Novice-High	Able to understand short, learned utterances and some sentence-length utterances, particularly where context strongly supports understanding and speech is clearly audible. Comprehends words and phrases from simple questions, statements, high-frequency commands, and courtesy formulae. May require repetition, rephrasing and/or a slowed rate of speech for comprehension.
Intermediate	The Intermediate level is characterized by an ability to understand main ideas and some facts from interactive exchanges and simple connected aural texts.
Intermediate-Low	Able to understand sentence-length utterances which consist of recombinations of learned elements in a limited number of content areas, particularly if strongly supported by the situational context. Content refers to basic personal background and needs, social conventions, and routine tasks, such as getting meals and receiving simple instructions and directions. Listening tasks pertain primarily to spontaneous face-to-face conversations. Understanding is often uneven; repetition and rewording may be necessary. Misunderstandings in both main ideas and details arise frequently.
Intermediate-Mid	Able to understand sentence-length utterances which consist of

	recombinations of learned utterances on a variety of topics. Content continues to refer primarily to basic personal background and needs, social conventions, and somewhat more complex tasks, such as lodging, transportation, and shopping. Additional content areas include some personal interests and activities, and a greater diversity of instructions and directions. Listening tasks not only pertain to spontaneous face-to-face conversations but also to short routine telephone conversations and some deliberate speech, such as simple announcements and reports over the media. Understanding continues to be uneven.
Intermediate-High	Able to sustain understanding over longer stretches of connected discourse on a number of topics pertaining to different times and places; however, understanding is inconsistent due to failure to grasp main ideas and/or details. Thus, while topics do not differ significantly from those of an Advanced-level listener, comprehension is less in quantity and poorer in quality.
Advanced	The Advanced level is characterized by an ability to understand main ideas and most details of connected discourse on a variety of topics beyond the immediacy of the situation, including some topics where comprehension is complicated due to an unexpected sequence of events.
Advanced	Able to understand main ideas and most details of connected discourse on a variety of topics beyond the immediacy of the situation. Comprehension may be uneven due to a variety of linguistic and extralinguistic factors, among which topic familiarity is very prominent. These texts frequently involve description and narration in different time frames or aspects, such as present, nonpast, habitual, or imperfective. Texts may include interviews, short lectures on familiar topics, and news items and reports primarily dealing with factual information. Listener is aware of cohesive devices but may not be able to use them to follow the sequence of thought in an oral text.
Advanced-Plus	Able to understand the main ideas of most speech in a standard dialect; however, the listener may not be able to sustain comprehension in extended discourse which is propositionally and linguistically complex. Listener shows an emerging awareness of culturally implied meanings beyond the surface meanings of the text but may fail to grasp sociocultural nuances of the message.
Superior	The Superior level is characterized by an ability to understand concrete and abstract topics in extended discourse offered by speakers using native-like discourse strategies.
Superior	Able to understand the main ideas of all speech in a standard dialect, including technical discussion in a field of specialization. Can follow the essentials of extended discourse which is propositionally and linguistically complex, as in academic/professional settings, in lectures, speeches, and reports. Listener shows some appreciation of aesthetic norms of target language, of idioms, colloquialisms, and register shifting. Able to make inferences within the cultural framework of the target language. Understanding is aided by an awareness of the underlying organizational structure of the oral text and includes sensitivity for its social and cultural references and its affective overtones. Rarely

misunderstands but may not understand excessively rapid, highly colloquial speech or speech that has strong cultural references.

Distinguished — The Distinguished level is characterized by an ability to understand accurately most linguistic styles and forms from within the cultural framework of the language.

Distinguished — Able to understand all forms and styles of speech pertinent to personal, social, and professional needs tailored to different audiences. Shows strong sensitivity to social and cultural references and aesthetic norms by processing language from within the cultural framework. Texts include theater plays, screen productions, editorials, symposia, academic debates, public policy statements, literary readings, and most jokes and puns. May have difficulty with some dialects and slang.

Generic Descriptions–Reading

These guidelines assume all reading texts to be authentic and legible.

Novice — The Novice level is characterized by an ability to
—identify isolated words and phrases when strongly supported by context; and
—identify learned material.

Novice-Low — Able occasionally to identify isolated words and/or major phrases when strongly supported by context.

Novice-Mid — Able to recognize the symbols of an alphabetic and/or syllabic writing system and/or a limited number of characters in a system that uses characters. The reader can identify an increasing number of highly contextualized words and/or phrases including cognates and borrowed words, where appropriate. Material understood rarely exceeds a single phrase at a time, and rereading may be required.

Novice-High — Has sufficient control of the writing system to interpret written language in areas of practical need. Where vocabulary has been learned, can read for instructional and directional purposes standardized messages, phrases, or expressions, such as some items on menus, schedules, timetables, maps, and signs. At times, but not on a consistent basis, the Novice-High-level reader may be able to derive meaning from material at a slightly higher level where context and/or extralinguistic background knowledge are supportive.

Intermediate — The Intermediate level is characterized by an ability to understand main ideas and some facts from simple connected texts.

Intermediate-Low — Able to understand main ideas and/or some facts from the simplest connected texts dealing with basic personal and social needs. Such texts are linguistically noncomplex and have a clear underlying internal structure, for example, chronological sequencing. They impart basic information about which the reader has to make only minimal suppositions or to which the reader brings personal interest and/or knowledge. Examples include messages with social purposes or information for the widest possible audience, such as public announce-

	ments and short, straightforward instructions dealing with public life. Some misunderstandings will occur.
Intermediate-Mid	Able to read consistently with increased understanding simple connected texts dealing with a variety of basic and social needs. Such texts are still linguistically noncomplex and have a clear underlying internal structure. They impart basic information about which the reader has to make minimal suppositions and to which the reader brings personal interest and/or knowledge. Examples may include short, straightforward descriptions of persons, places, and things written for a wide audience.
Intermediate-High	Able to read consistently with full understanding simple connected texts dealing with basic personal and social needs about which the reader has personal interest and/or knowledge. Can get some main ideas and information from texts at the next higher level featuring description and narration. Structural complexity may interfere with comprehension; for example, basic grammatical relations may be misinterpreted and temporal references may rely primarily on lexical items. Has some difficulty with the cohesive factors in discourse, such as matching pronouns with referents. While texts do not differ significantly from those at the Advanced level, comprehension is less consistent. May have to read material several times for understanding.
Advanced	The Advanced level is characterized by —an ability to read with consistent understanding prose several paragraphs in length, dealing primarily with factual information and intended for the general reader; and —in areas of special interest or knowledge, an increasing ability to understand parts of texts which are propositionally and linguistically complex.
Advanced	Able to read somewhat longer prose of several paragraphs in length, particularly if presented with a clear underlying structure. The prose is predominantly in familiar sentence patterns. Reader gets the main ideas and facts and misses some details. Comprehension derives not only from situational and subject-matter knowledge but also from increasing control of the language. Texts at this level include descriptions and narrations such as simple short stories, news items, bibliographical information, social notices, personal correspondence, routinized business letters, and simple technical material written for the general reader.
Advanced-Plus	Able to follow essential points of written discourse at the Superior level in areas of special interest or knowledge. Able to understand parts of texts which are conceptually abstract and linguistically complex, and/or texts which treat unfamiliar topics and situations, as well as some texts which involve aspects of target-language culture. Able to comprehend the facts to make appropriate inferences. An emerging awareness of the aesthetic properties of language and of its literary styles permits comprehension of a wider variety of texts, including literary. Misunderstandings may occur.
Superior	The Superior level is characterized by an ability to read, for information or for pleasure with almost complete comprehension and at normal speed, a wide variety of texts on a wide variety of topics.

Superior	Able to read with almost complete comprehension and at normal speed expository prose on unfamiliar subjects and a variety of literary texts. Reading ability is not dependent on subject-matter knowledge, although the reader is not expected to comprehend thoroughly texts which are highly dependent on knowledge of the target culture. Reads easily for pleasure. Superior-level texts feature hypotheses, argumentation, and supported opinions and include grammatical patterns and vocabulary ordinarily encountered in academic/professional reading. At this level, due to the control of general vocabulary and structure, the reader is almost always able to match the meanings derived from extralinguistic knowledge with meanings derived from knowledge of the language, allowing for smooth and efficient reading of diverse texts. Occasional misunderstandings may still occur; for example, the reader may experience some difficulty with unusually complex structures and low-frequency idioms. At the Superior level the reader can match strategies, top-down or bottom-up, which are most appropriate to the text. (Top-down strategies rely on real-world knowledge and prediction based on genre and organizational scheme of the text. Bottom-up strategies rely on actual linguistic knowledge.) Material at this level will include a variety of literary texts, editorials, correspondence, general reports, and technical material in professional fields. Rereading is rarely necessary, and misreading is rare.
Distinguished	The Distinguished level is characterized by an ability to read fluently and accurately most styles and forms of the language with comprehension that is achieved from within the cultural framework of the language and that includes appreciation of nuance and subtlety.
Distinguished	Able to read fluently and accurately most styles and forms of the language pertinent to academic and professional needs. Able to relate inferences in the text to real-world knowledge and understand almost all sociolinguistic and cultural references by processing language from within the cultural framework. Able to understand a writer's use of nuance and subtlety. Can readily follow unpredictable turns of thought and author intent in such materials as sophisticated editorials, specialized journal articles, and literary texts such as novels, plays, poems, as well as in any subject-matter area directed to the general reader.

Generic Descriptions–Writing

Novice	The Novice level is characterized by an ability to produce isolated words and phrases.
Novice-Low	Able to form some letters in an alphabetic system. In languages whose writing systems use syllabaries or characters, writer is able to both copy and produce the basic strokes. Can produce romanization of isolated characters, where applicable.
Novice-Mid	Able to copy or transcribe familiar words or phrases and reproduce some from memory. No practical communicative writing skills.
Novice-High	Able to write simple fixed expressions and limited memorized material and some recombinations thereof. Can supply information on sim-

ple forms and documents. Can write names, numbers, dates, own nationality, and other simple autobiographical information, as well as some short phrases and simple lists. Can write all the symbols in an alphabetic or syllabic system or 50–100 characters or compounds in a character writing system. Spelling and representation of symbols (letters, syllables, characters) may be partially correct.

Intermediate — The Intermediate level is characterized by an ability to meet practical writing needs by communicating simple facts and ideas in a loose collection of sentences.

Intermediate-Low — Able to meet limited practical writing needs. Can write short messages, postcards, and take down simple notes, such as telephone messages. Can create statements or questions within the scope of limited language experience. Material produced consists of recombinations of learned vocabulary and structures into simple sentences on very familiar topics. Language is inadequate to express in writing anything but elementary needs. Frequent errors in grammar, vocabulary, punctuation, spelling, and in formation of nonalphabetic symbols, but writing can be understood by natives used to the writing of nonnatives.

Intermediate-Mid — Able to meet a number of practical writing needs. Can write short, simple letters. Content involves personal preferences, daily routine, everyday events, and other topics grounded in personal experience. Can express present time or at least one other time frame or aspect consistently, e.g., nonpast, habitual, imperfective. Evidence of control of the syntax of noncomplex sentences and basic inflectional morphology, such as declensions and conjugation. Writing tends to be a loose collection of sentences or sentence fragments on a given topic and provides little evidence of conscious organization. Can be understood by natives used to the writing of nonnatives.

Intermediate-High — Able to meet most practical writing needs and limited social demands. Can take notes in some detail on familiar topics and respond in writing to personal questions. Can write simple letters, brief synopses and paraphrases, summaries of biographical data, work and school experience. In those languages relying primarily on content words and time expressions to express time, tense, or aspect, some precision is displayed; where tense and/or aspect is expressed through verbal inflection, forms are produced rather consistently, but not always accurately. An ability to describe and narrate in paragraphs is emerging. Rarely uses basic cohesive elements, such as pronominal substitutions or synonyms in written discourse. Writing, though faulty, is generally comprehensible to natives used to the writing of nonnatives.

Advanced — The Advanced level is characterized by an ability to write narratives and descriptions of a factual nature of at least several paragraphs in length on familiar topics.

Advanced — Able to write routine social correspondence and join sentences in simple discourse of at least several paragraphs in length on familiar topics. Can write simple social correspondence, take notes, write cohesive summaries and resumes, as well as narratives and descriptions of a factual nature. Has sufficient writing vocabulary to express self simply with some circumlocution. May still make errors in punctuation, spelling, or the formation of nonalphabetic symbols. Good control of the

	morphology and the most frequently used syntactic structures, e.g., common word order patterns, coordination, subordination, but makes frequent errors in producing complex sentences. Uses a limited number of cohesive devices, such as pronouns, accurately. Writing may resemble literal translations from the native language, but a sense of organization (rhetorical structure) is emerging. Writing is understandable to natives not used to the writing of nonnatives.
Advanced-Plus	Able to write about a variety of topics with significant precision and in detail. Can write most social and informal business correspondence. Can describe and narrate personal experiences fully but has difficulty supporting points of view in written discourse. Can write about the concrete aspects of topics relating to particular interests and special fields of competence. Often shows remarkable fluency and ease of expression, but under time constraints and pressure writing may be inaccurate. Generally strong in either grammar or vocabulary, but not in both. Weakness and unevenness in one of the foregoing or in spelling or character writing formation may result in occasional miscommunication. Some misuse of vocabulary may still be evident. Style may still be obviously foreign.
Superior	The Superior level is characterized by an ability to write formally and informally on practical, social, and professional topics.
Superior	Able to express self effectively in most formal and informal writing on practical, social, and professional topics. Can write most types of correspondence, such as memos as well as social and business letters, and short research papers and statements of position in areas of special interest or in special fields. Good control of a full range of structures, spelling or nonalphabetic symbol production, and a wide general vocabulary allow the writer to hypothesize and present arguments or points of view accurately and effectively. An underlying organization, such as chronological ordering, logical ordering, cause and effect, comparison, and thematic development is strongly evident, although not thoroughly executed and/or not totally reflecting target-language patterns. Although sensitive to differences in formal and informal style, still may not tailor writing precisely to a variety of purposes and/or readers. Errors in writing rarely disturb natives or cause miscommunication.

Appendix 2
ILR Language Skill Level Descriptions *

INTERAGENCY LANGUAGE ROUND TABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
LISTENING

Preface

The following proficiency level descriptions characterize comprehension of the spoken language. Each of the six "base levels" (Coded 00, 10, 20, 30, 40, and 50) implies control of any previous "base level" functions and accuracy. The "plus level" designation (coded 06, 16, 26, 36, and 46) will be assigned when the language skill substantially exceeds one base skill level and does not fully meet the criteria for the next base level. Therefore, the plus level descriptions are supplementary to the base level descriptions.

A skill level is assigned to a person through an authorized language examination. Examiners assign a level on a variety of performance criteria exemplified in the descriptive statements. Therefore, the examples given here illustrate, but do not exhaustively describe, either the skills a person may possess or situations in which he/she may function effectively.

Statements describing accuracy refer to typical stages in the development of competence in the most commonly taught languages in formal instructional programs. In other languages, emerging competence parallels these characterizations, but often with different details.

Unless otherwise specified, the term "native speaker" refers to native speakers of a standard dialect.

"Well-educated," in the context of these skills descriptions, does not necessarily imply formal higher education. However, in cultures where formal higher education is common, the language-use abilities of persons who have had such education is considered the standard. In other words, such a person meets the contemporary expectations for the formal, careful style of the language, as well as a range of less formal varieties of the language.

Listening 0 (No Proficiency)

No practical understanding of the spoken language. Understanding is limited to occasional isolated words with essentially no ability to comprehend communication. (Has been coded L-0 in some nonautomated applications.) (Data Code 00)

Listening 0+ (Memorized Proficiency)

Sufficient comprehension to understand a number of memorized utterances in areas of immediate needs. Slight increase in utterance length is understood but requires frequent long pauses between understood phrases and repeated requests on the listener's part for repetition. Understands with reasonable accuracy only when they involve short memorized utterances or formulae. Utterances understood are relatively short in length. Misunderstandings arise due to ignoring or inaccurately hearing sounds or word endings (both inflectional and non-inflectional), thus distorting the original meaning. Can understand only with difficulty even persons such as teachers who are used to speaking with non-native speakers. Can best understand those statements where context strongly supports the utterance meaning. Gets some main ideas. (Has been coded L-0+ in some nonautomated applications.) (Data Code 06)

Listening 1 (Elementary Proficiency)

Sufficient comprehension to understand utterances about basic survival needs and minimum courtesy and travel requirements. In areas of immediate need or on very familiar topics, can understand simple questions and answers, simple statements, and very simple face-to-face conversations in a standard dialect. These must often be delivered more clearly than normal at a rate slower than normal, with frequent repetitions or paraphrases (that is, by a native used to dealing with foreigners). Once learned, these sentences can be varied for similar level vocabulary and grammar and still be understood. In the majority of utterances, misunderstandings arise due to overlooked or misunderstood syntax and other grammatical clues. Comprehension vocabulary is inadequate to understand anything but the most elementary needs. Strong interference from the candidate's native language occurs. Little precision in the information is understood owing to tentative state of passive grammar and lack of vocabulary. Comprehension areas include

* *Defense Language Institute. (1992). Defense Language Proficiency Test IV: Familiarization Guide (DLIFLC Pamphlet 350-14, pp. A-1 to A-11). Monterey, CA: Author.*

BEST COPY AVAILABLE

basic needs such as: meals, lodging, transportation, time, and simple directions (including both route instructions and orders from customs officials, policemen, etc.). Understands main ideas. (Has been coded L-1 in some nonautomated applications.) (Data Code 10)

Listening 1 + (Elementary Proficiency, Plus)

Sufficient comprehension to understand short conversations about all survival needs and limited social demands. Developing flexibility evident in understanding into a range of circumstances beyond immediate survival needs. Shows spontaneity in understanding by speed, although consistency of understanding uneven. Limited vocabulary range necessitates repetition for understanding. Understands commoner time forms and most question forms, some word order patterns, but miscommunication still occurs with more complex patterns. Cannot sustain understanding of coherent structures in longer utterances or in unfamiliar situations. Understanding of descriptions and the giving of precise information is limited. Aware of basic cohesive features, e.g., pronouns, verb inflections, but many are unreliably understood, especially if less immediate in reference. Understanding is largely limited to a series of short, discrete utterances. Still has to ask for utterances to be repeated. Some ability to understand facts. (Has been coded L-1+ in some nonautomated applications.) (Data Code 16)

Listening 2 (Limited Working Proficiency)

Sufficient comprehension to understand conversations on routine social demands and limited job requirements. Able to understand face-to-face speech in a standard dialect, delivered at a normal speaking rate with some repetition and rewording, by a native speaker not used to dealing with foreigners, about everyday topics, common personal and family news, well-known current events, and routine office matters through descriptions and narration about current, past, and future events; can follow essential points of discussion or speech at an elementary level on topics in his/her special professional field. Only understands occasional words and phrases of statements made in unfavorable conditions, for example through loudspeakers outdoors. Understands factual content. Native language causes less interference in listening comprehension. Able to understand facts, i.e., the lines, but not between or beyond the lines. (Has been coded L-2 in some nonautomated applications.) (Data Code 20)

Listening 2 + (Limited Working Proficiency, Plus)

Sufficient comprehension to understand most routine social demands and most conversations on work requirements, as well as some discussions on concrete topics related to particular interest and special fields of competence. Often shows remarkable ability and ease of understanding but may breakdown under tension or pressure. Candidate may display weakness or deficiency due to inadequate vocabulary base or less than secure knowledge of grammar and syntax. Normally understands general vocabulary with some hesitant understanding of everyday vocabulary still evident. Can sometimes detect emotional overtones. Some ability to understand implications. (Has been coded L-2+ in some nonautomated applications.) (Data Code 26)

Listening 3 (General Professional Proficiency)

Able to understand the essentials of all speech in a standard dialect, including technical discussions within a special field. Has effective understanding of face-to-face speech, delivered with normal clarity and speed in a standard dialect, on general topics and areas of special interest; understands hypothesizing and supported opinions. Has broad enough vocabulary that rarely has to ask for paraphrasing or explanation. Can follow accurately the essentials in conversations between educated native speakers, reasonably clear telephone calls, radio broadcasts, news stories similar to wire service reports, oral reports, some oral technical reports, and public addresses on non-technical subjects; can understand without difficulty all forms of standard speech concerning a special professional field. Does not understand native speakers if they speak very quickly or use slang or dialect. Can often detect emotional overtones. Can understand implications. (Has been coded L-3 in some nonautomated applications.) (Data Code 30)

Listening 3 + (General Professional Proficiency, Plus)

BEST COPY AVAILABLE

Comprehends most of the content and intent of a variety of forms and styles of speech pertinent to professional needs, as well as general topics and social conversation. Ability to comprehend many socio-linguistic and cultural references. However, may miss some subtleties and nuances.

Increased ability to comprehend unusually complex structures in lengthy utterances and to comprehend many distinctions in language tailored for different audiences. Increased ability to understand native speakers talking quickly, using nonstandard dialect or slang; however, comprehension not complete. Can discern some relationships among sophisticated listening materials in the context of broad experience. Can follow some unpredictable turns of thought readily, for example, in informal and formal speeches covering editorial, conjectural, and literary materials in subject matter areas directed to the general listener. (Has been coded L-3+ in some nonautomated applications.) (Data Code 36)

Listening 4 (Advanced Professional Proficiency)

Able to understand all forms and styles of speech pertinent to professional needs. Able to understand fully all of speech with extensive and precise vocabulary, subtleties, and nuances in all standard dialects on any subject relevant to professional needs within the range of his/her experience, including social conversations; all intelligible broadcasts and telephone calls; and many kinds of technical discussions and discourse. Understands language specifically tailored (including persuasion, representation, counseling, and negotiating) to different audiences. Able to understand the essentials of speech in some nonstandard dialects. Has difficulty in understanding extreme dialect and slang, also in understanding speech in unfavorable conditions, for example through bad loudspeakers outdoors. Can discern relationships among sophisticated listening materials in the context of broad experience. Can follow unpredictable turns of thought readily, for example, in informal and formal speeches covering editorial, conjectural, and literary materials in any subject matter directed to the general listener. (Has been coded L-4 in some nonautomated applications.) (Data Code 40)

Listening 4+ (Advanced Professional Proficiency, Plus)

Increased ability to understand extremely difficult and abstract speech, as well as ability to understand all forms and styles of speech pertinent to professional needs, including social conversations. Increased ability to comprehend native speakers using extreme nonstandard dialects and slang, as well as to understand speech in unfavorable conditions. Strong sensitivity to socio-linguistic and cultural references. Accuracy is close to that of the well-educated native listener but still not equivalent. (Has been coded L-4+ in some nonautomated applications.) (Data Code 46)

Listening 5 (Functionally Native Proficiency)

Comprehension equivalent to that of the well-educated native listener. Able to understand fully all forms and styles of speech intelligible to the well-educated native listener, including a number of regional and illiterate dialects, highly .olloquial speech and conversations, and discourse distorted by marked interference from other noise. Able to understand how natives think as they create discourse. Able to understand extremely difficult and abstract speech. (Has been coded L-5 in some nonautomated applications.) (Data Code 50)

BEST COPY AVAILABLE

INTERAGENCY LANGUAGE ROUND TABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
READING

Preface

The following proficiency level descriptions characterize comprehension of the written language. Each of the six "base levels" (Coded 00, 10, 20, 30, 40, and 50) implies control of any previous "base level" functions and accuracy. The "plus level" designation (coded 06, 16, 26, 36, and 46) will be assigned when the language skill substantially exceeds one base skill level and does not fully meet the criteria for the next base level. Therefore, the plus level descriptions are supplementary to the base level descriptions.

A skill level is assigned to a person through an authorized language examination. Examiners assign a level on a variety of performance criteria exemplified in the descriptive statements. Therefore, the examples given here illustrate, but do not exhaustively describe, either the skills a person may possess or situations in which he/she may function effectively.

Statements describing accuracy refer to typical stages in the development of competence in the most commonly taught languages in formal instructional programs. In other languages, emerging competence parallels these characterizations, but often with different details.

Unless otherwise specified, the term "native speaker" refers to native speakers of a standard dialect.

"Well-educated," in the context of these skills descriptions, does not necessarily imply formal higher education. However, in cultures where formal higher education is common, the language-use abilities of persons who have had such education is considered the standard. In other words, such a person meets the contemporary expectations for the formal, careful style of the language, as well as a range of less formal varieties of the language.

Reading 0 (No Proficiency)

No practical ability to read the language. Consistently misunderstands or cannot comprehend the written language. (Has been coded R-0 in some nonautomated applications.) (Data Code 00)

Reading 0+ (Memorized Proficiency)

Can recognize all the letters in the printed version of an alphabetic system and high-frequency elements of a syllabary or character system. Able to read some or all of the following: numbers, isolated words and phrases, personal and place names, street signs, office and shop designations. The above often interpreted inaccurately. (Has been coded R-0+ in some nonautomated applications.) (Data Code 06)

Reading 1 (Elementary Proficiency)

Sufficient comprehension to read very simple connected written material in a form equivalent to usual printing or typescript. Can read either representations of familiar formulaic verbal exchanges or simple language containing only the highest frequency structural patterns and vocabulary, including shared international vocabulary items and cognates (when appropriate). Able to read and understand known language elements that have been recombined in new ways to achieve different meanings at a similar level of simplicity. Texts may include simple narratives of routine behavior; highly predictable descriptions of persons, places, or things; and explanations of geography and government such as those simplified for tourists. Some misunderstandings possible on simple texts. Can get some main ideas and locate prominent items of professional significance in more complex texts. Can identify general subject matter in some authentic texts. (Has been coded R-1 in some nonautomated applications.) (Data Code 10)

Reading 1+ (Elementary Proficiency, Plus)

Sufficient comprehension to understand simple discourse in printed form for informative social purposes. Can read material such as announcements of public events, simple prose

containing biographical information or narration of events, and straightforward newspaper headlines. Can guess at unfamiliar vocabulary if highly contextualized, but with difficulty in unfamiliar contexts. Can get some main ideas and locate routine information of professional significance in more complex texts. Can follow essential points of written discussion at an elementary level in topics in his/her special professional field. In commonly taught languages, the individual may not control the structure well. For example, basic grammatical relations are often misinterpreted, and temporal reference may rely primarily on lexical items as time indicators. Has some difficulty with the cohesive factors in the discourse such as matching pronouns with referents. May have to read materials several times for understanding. (Has been coded R-1+ in some nonautomated applications.) (Data Code 16)

Reading 2 (Limited Working Proficiency)

Sufficient comprehension to read simple, authentic written material in a form equivalent to usual printing or typescript on subjects within a familiar context. Able to read straightforward, familiar, factual material with some misunderstandings, but in general insufficiently experienced with the language to draw inferences directly from the linguistic aspects of the text. Can locate and understand the main ideas and details in material written for the general reader. However, persons who have professional knowledge of a subject may be able to summarize or perform sorting and locating tasks with written texts that are well beyond their general proficiency level. The individual can read uncomplicated, but authentic prose on familiar subjects that are normally presented in a predictable sequence which aids the reader in understanding. Texts may include descriptions and narrations in contexts such as news items describing frequently occurring events, simple biographical information, social notices, formulaic business letters, and simple technical materials written for the general reader. Generally, the prose that can be read by the individual is predominantly in straightforward / high-frequency sentence patterns. The individual does not have a broad active vocabulary (that is, which he/she recognizes immediately on sight) but is able to use contextual and real-world cues to understand the text. Characteristically, however, the individual is quite slow in performing such a process. He/she is typically able to answer factual questions about authentic texts of the types described above. (Has been coded R-2 in some nonautomated applications.) (Data Code 20)

Reading 2+ (Limited Working Proficiency, Plus)

Sufficient comprehension to understand most factual material in nontechnical prose, as well as some discussions on concrete topics related to special professional interests. Is markedly more proficient at reading materials on familiar topics. Is able to separate the main ideas and details from lesser ones and uses that distinction to advance comprehension. The individual is able to use linguistic context and real-world knowledge to make sensible guesses about unfamiliar material. Has a broad active reading vocabulary. The individual is able to get the gist of main and subsidiary ideas in texts which could be read thoroughly only by persons with much higher proficiencies. Weaknesses include slowness, uncertainty, and inability to discern nuances and/or intentionally disguised meaning. (Has been coded R-2+ in some nonautomated applications.) (Data Code 26)

Reading 3 (General Professional Proficiency)

Able to read within range of speed and with almost complete comprehension a variety of authentic prose material on unfamiliar subjects. Reading ability is not dependent on subject matter knowledge, although it is not expected that the individual can comprehend thoroughly subject matter which is highly dependent on cultural knowledge or which is outside his/her general experience and not accompanied by explanation. Text-types include news stories similar to wire service reports or internal news items in major periodicals, routine correspondence, general reports, and technical material in his/her professional field; all of these may include hypothesis, argumentation, and supported opinions. Misreading rare. Almost always able to interpret material correctly, relate ideas, and "read between the lines," (that is, understand the writers' implicit intents in texts of the above types). Can get the gist of more sophisticated texts but may be unable to detect or understand subtlety and nuance. Rarely has to pause over or reread general vocabulary.

BEST COPY AVAILABLE

However, may experience some difficulty with unusually complex structure and low-frequency idioms. (Has been coded R-3 in some nonautomated applications.) (Data Code 30)

Reading 3+ (General Professional Proficiency, Plus)

Can comprehend a variety of styles and forms pertinent to professional needs. Rarely misinterprets such texts or rarely experiences difficulty relating ideas or making inferences. Able to comprehend many socio-linguistic and cultural references. However, may miss some nuances and subtleties. Able to comprehend a considerable range of intentionally complex structures, low-frequency idioms, and uncommon connotative intentions; however, accuracy is not complete. The individual is typically able to read with facility, understand, and appreciate contemporary expository, technical, or literary texts which do not rely heavily on slang and unusual idioms. (Has been coded R-3+ in some nonautomated applications.) (Data Code 36)

Reading 4 (Advanced Professional Proficiency)

Able to read fluently and accurately all styles and forms of the language pertinent to professional needs. The individual's experience with the written language is extensive enough that he/she is able to relate inferences in the text to real-world knowledge and understand almost all socio-linguistic and cultural references. Able to "read beyond the lines" (that is, understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment). Able to read and understand the intent of writers' employment of nuance and subtlety. The individual can discern relationships among sophisticated written materials in the context of broad experience. Can follow unpredictable turns of thought readily, for example, in editorials, conjectural, and literary texts in any subject matter area directed to the general reader. Can read essentially all materials in his/her special field, including official and professional documents and correspondence. Recognizes all professionally relevant vocabulary known to the educated nonprofessional native, although may have some difficulty with slang. Can read reasonably legible handwriting without difficulty. Accuracy is often nearly that of a well-educated native reader. (Has been coded R-4 in some nonautomated applications.) (Data Code 40)

Reading 4+ (Advanced Professional Proficiency, Plus)

Near native ability to read and understand extremely difficult or abstract prose, a very wide variety of vocabulary, idioms, colloquialisms, and slang. Strong sensitivity to and understanding of socio-linguistic and cultural references. Little difficulty in reading less than fully legible handwriting. Broad ability to "read beyond the lines" (that is, understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment) is nearly that of a well-read or well-educated native reader. Accuracy is close to that of the well-educated native reader, but not equivalent. (Has been coded R-4+ in some nonautomated applications.) (Data Code 46)

Reading 5 (Functionally Native Proficiency)

Reading proficiency is functionally equivalent to that of the well-educated native reader. Can read extremely difficult and abstract prose; for example, general legal and technical as well as highly colloquial writings. Able to read literary texts, typically including contemporary avant-garde prose, poetry, and theatrical writing. Can read classical/archaic forms of literature with the same degree of facility as the well-educated, but nonspecialist native. Reads and understands a wide variety of vocabulary and idioms, colloquialisms, slang, and pertinent cultural references. With varying degrees of difficulty, can read all kinds of handwritten documents. (Has been coded R-5 in some nonautomated applications.) (Data Code 50)

INTERAGENCY LANGUAGE ROUND TABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
SPEAKING

Preface

The following proficiency level descriptions characterize spoken language use. Each of the six "base levels" (Coded 00, 10, 20, 30, 40, and 50) implies control of any previous "base level" functions and accuracy. The "plus level" designation (coded 06, 16, 26, 36, and 46) will be assigned when the language skill substantially exceeds one base skill level and does not fully meet the criteria for the next base level. Therefore, the plus level descriptions are supplementary to the base level descriptions.

A skill level is assigned to a person through an authorized language examination. Examiners assign a level on a variety of performance criteria exemplified in the descriptive statements. Therefore, the examples given here illustrate, but do not exhaustively describe, either the skills a person may possess or situations in which he/she may function effectively.

Statements describing accuracy refer to typical stages in the development of competence in the most commonly taught languages in formal instructional programs. In other languages, emerging competence parallels these characterizations, but often with different details.

Unless otherwise specified, the term "native speaker" refers to native speakers of a standard dialect.

"Well-educated," in the context of these skills descriptions, does not necessarily imply formal higher education. However, in cultures where formal higher education is common, the language-use abilities of persons who have had such education is considered the standard. In other words, such a person meets the contemporary expectations for the formal, careful style of the language, as well as a range of less formal varieties of the language.

Speaking 0 (No Proficiency)

Unable to function in the spoken language. Oral production is limited to occasional isolated words. Has essentially no communicative ability. (Has been coded S-0 in some nonautomated applications.) (Data Code 00)

Speaking 0 + (Memorized Proficiency)

Able to satisfy immediate needs using rehearsed utterances. Shows little real autonomy of expression, flexibility, or spontaneity. Can ask questions or make statements with reasonable accuracy only with memorized utterances or formulas. Attempts at creating speech are unsuccessful.

Examples: The individual's vocabulary is usually limited to areas of immediate survival needs. Most utterances are telegraphic; that is, functors (linking words, markers, and the like) are omitted, confused, or distorted. An individual can usually differentiate most significant sounds when produced in isolation, but, when combined in words or groups of words, errors may be frequent. Even with repetition, communication is severely limited even with persons used to dealing with foreigners. Stress, intonation, tone, etc., are usually quite faulty. (Has been coded S-0+ in some nonautomated applications.) (Data Code 06)

Speaking 1 (Elementary Proficiency)

Able to satisfy minimum courtesy requirements and maintain very simple face-to-face conversations on familiar topics. A native speaker must often use slowed speech, repetition, paraphrase, or a combination of these to be understood by this individual. Similarly, the native speaker must strain and employ real-world knowledge to understand even simple statements/ questions from this individual. This speaker has a functional, but limited proficiency. Misunderstandings are frequent, but the individual is able to ask for help and to verify comprehension of native speech in face-to-face interaction. The individual is unable to produce continuous discourse except with rehearsed material.

Examples: Structural accuracy is likely to be random or severely limited. Time concepts are vague. Vocabulary is inaccurate, and its range is

very narrow. The individual often speaks with great difficulty. By repeating, such speakers can make themselves understood to native speakers who are in regular contact with foreigners, but there is little precision in the information conveyed. Needs, experience, or training may vary greatly from individual to individual; for example, speakers at this level may have encountered quite different vocabulary areas. However, the individual can typically satisfy predictable, simple, personal, and accommodation needs; can generally meet courtesy, introduction, and identification requirements; exchange greetings; and elicit and provide, for example, predictable and skeletal biographical information. He/she might give information about business hours, explain routine procedures in a limited way, and state in a simple manner what actions will be taken. He/she is able to formulate some questions even in languages with complicated question constructions. Almost every utterance may be characterized by structural errors and errors in basic grammatical relations. Vocabulary is extremely limited and characteristically does not include modifiers. Pronunciation, stress, and intonation are generally poor, often heavily influenced by another language. Use of structure and vocabulary is highly imprecise. (Has been coded S-1 in some nonautomated applications.) (Data Code 10)

Speaking 1+ (Elementary Proficiency, Plus)

Can initiate and maintain predictable face-to-face conversations and satisfy limited social demands. He/she may, however, have little understanding of the social conventions of conversation. The interlocutor is generally required to strain and employ real-world knowledge to understand even some simple speech. The speaker at this level may hesitate and may have to change subjects due to lack of language resources. Range and control of the language are limited. Speech largely consists of a series of short, discrete utterances.

Examples: The individual is able to satisfy most travel and accommodation needs and a limited range of social demands beyond exchange of skeletal biographic information. Speaking ability may extend beyond immediate survival needs. Accuracy in basic grammatical relations is evident, although not consistent. May exhibit the more common forms of verb tenses, for example, but may make frequent errors in formation and selection. While some structures are established, errors occur in more complex patterns. The individual typically cannot sustain coherent structures in longer utterances or unfamiliar situations. Ability to describe and give precise information is limited. Person, space, and time references are often used incorrectly. Pronunciation is understandable to natives used to dealing with foreigners. Can combine most significant sounds with reasonable comprehensibility but has difficulty in producing certain sounds in certain positions or in certain combinations. Speech will usually be labored. Frequently has to repeat utterances to be understood by the general public. (Has been coded S-1+ in some nonautomated applications.) (Data Code 16)

Speaking 2 (Limited Working Proficiency)

Able to satisfy routine social demands and limited work requirements. Can handle routine work-related interactions that are limited in scope. In more complex and sophisticated work-related tasks, language usage generally disturbs the native speaker. Can handle with confidence, but not facility, most normal high-frequency social conversational situations, including extensive but casual conversations about current events, as well as work, family, and autobiographical information. The individual can get the gist of most everyday conversations but has some difficulty understanding native speakers in situations that require specialized or sophisticated knowledge. The individual's utterances are minimally cohesive. Linguistic structure is usually not very elaborate and not thoroughly controlled; errors are frequent. Vocabulary use is appropriate for high-frequency utterances, but unusual or imprecise elsewhere.

Examples: While these interactions will vary widely from individual to individual, the individual can typically ask and answer predictable questions in the workplace and give straightforward instructions to subordinates. Additionally, the individual can participate in personal and accommodation type interactions with elaboration and facility; that is, can give and understand complicated, detailed, and extensive directions and make nonroutine changes in travel and accommodation arrangements. Simple structures and basic grammatical relations are typically controlled. How-

ever, there are areas of weakness. For example, in the commonly taught languages, these may be simple markings such as plurals, articles, linking words, and negatives or more complex structures such as tense/aspect usage, case morphology, passive constructions, word order, and embedding. (Has been coded S-2 in some nonautomated applications.) (Data Code 20)

Speaking 2+ (Limited Working Proficiency, Plus)

Able to satisfy most work requirements with language usage that is often, but not always, acceptable and effective. The individual shows considerable ability to communicate effectively on topics relating to particular interests and special fields of competence. Often shows a high degree of fluency and ease of speech, yet when under tension or pressure, the ability to use the language effectively may deteriorate. Comprehension of normal native speech is typically nearly complete. The individual may miss cultural and local references and may require a native speaker to adjust to his/her limitations in some ways. Native speakers often perceive the individual's speech to contain awkward or inaccurate phrasing of ideas, mistaken time, space, and person references, or to be inappropriate in some way, if not strictly incorrect.

Examples: Typically, the individual can participate in most social, formal, and informal interactions; but limitations either in range of contexts, types of tasks, or level of accuracy hinder effectiveness. The individual may be ill at ease with the use of the language either in social interaction or in speaking at length in professional contexts. He/ she is generally strong in either structural precision or vocabulary, but not in both. Weakness or unevenness in one of the foregoing, or in pronunciation, occasionally results in miscommunication. Normally controls, but cannot always easily produce, general vocabulary. Discourse is often incohesive. (Has been coded S-2+ in some nonautomated applications.) (Data Code 26)

Speaking 3 (General Professional Proficiency)

Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Nevertheless, the individual's limitations generally restrict the professional contexts of language use to matters of shared knowledge and/or international convention. Discourse is cohesive. The individual uses the language acceptably, but with some noticeable imperfections; yet, errors virtually never interfere with understanding and rarely disturb the native speaker. The individual can effectively combine structure and vocabulary to convey his/her meaning accurately. The individual speaks readily and fills pauses suitably. In face-to-face conversations with natives speaking the standard dialect at a normal rate of speech, comprehension is quite complete. Although cultural references, proverbs, and the implications of nuances and idioms may not be fully understood, the individual can easily repair the conversation. Pronunciation may be obviously foreign. Individual sounds are accurate; but stress, intonation, and pitch control may be faulty.

Examples: Can typically discuss particular interests and special fields of competence with reasonable ease. Can use the language as part of normal professional duties such as clarifying points, answering objections, justifying decisions, understanding the essence of challenges, stating and defending policy, conducting meetings, delivering briefings, or other extended and elaborate informative monologues. Can reliably elicit information and informed opinion from native speakers. Structural inaccuracy is rarely the major cause of misunderstanding. Use of structural devices is flexible and elaborate. Without searching for words or phrases, the individual uses the language clearly and relatively naturally to elaborate concepts freely and make ideas easily understandable to native speakers. Errors occur in low-frequency and highly complex structures. (Has been coded S-3 in some nonautomated applications.) (Data Code 30)

Speaking 3+ (General Professional Proficiency, Plus)

Is often able to use the language to satisfy professional needs in a wide rage of sophisticated and demanding tasks.

Examples: Despite obvious strengths, may exhibit some hesitancy, uncertainty, effort, or errors which limit the range of language-use tasks that can be reliably performed. Typically,

there is particular strength in fluency and one or more, but not all, of the following: breadth of lexicon, including items of low- and medium-frequencies, especially sociolinguistic/cultural references and nuances of close synonyms; structural precision, with sophisticated features that are readily, accurately, and appropriately controlled (such as complex modification and embedding in Indo-European languages); discourse competence in a wide range of contexts and tasks, often matching a native speaker's strategic and organizational abilities and expectations. Occasional patterned errors occur in low-frequency and highly complex structures. (Has been coded S-3+ in some nonautomated applications.) (Data Code 36)

Speaking 4 (Advanced Professional Proficiency)

Able to use the language fluently and accurately on all levels normally pertinent to professional needs. The individual's language usage and ability to function are fully successful. Organizes discourse well, using appropriate rhetorical speech devices, native cultural references, and understanding. Language ability only rarely hinders him/her in performing any task requiring language; yet, the individual would seldom be perceived as a native. Speaks effortlessly and smoothly and is able to use the language with a high degree of effectiveness, reliability, and precision for all representational purposes within the range of personal and professional experience and scope of responsibilities. Can serve as an informal interpreter in a range of unpredictable circumstances. Can perform extensive, sophisticated language tasks, encompassing most matters of interest to well-educated native speakers, including tasks which do not bear directly on a professional specialty.

Examples: Can discuss in detail concepts which are fundamentally different from those of the target culture and make those concepts clear and accessible to the native speaker. Similarly, the individual can understand the details and ramifications of concepts that are culturally or conceptually different from his/her own. Can set the tone of interpersonal official, semi-official, and nonprofessional verbal exchanges with a representative range of native speakers (in a range of varied audiences, purposes, tasks, and settings). Can play an effective role among native speakers in such contexts as conferences, nonstigmatized dialect. (Has been coded S-5 in lectures, and debates on matters of disagreement. Can advocate a position at length, both formally and in chance encounters, using sophisticated verbal strategies. Can understand and reliably produce shifts of both subject matter and tone. Can understand native speakers of the standard and other major dialects in essentially any face-to-face interaction. (Has been coded S-4 in some nonautomated applications.) (Data Code 40)

Speaking 4+ (Advanced Professional Proficiency, Plus)

Speaking proficiency is regularly superior in all respects, usually equivalent to that of a well-educated, highly articulate native speaker. Language ability does not impede the performance of any language-use task. However, the individual would not necessarily be perceived as culturally native.

Examples: The individual organizes discourse well, employing functional rhetorical speech devices, native cultural references and understanding. Effectively applies a native speaker's social and circumstantial knowledge. However, cannot sustain that performance under all circumstances. While the individual has a wide range and control of structure, an occasional non-native slip may occur. The individual has a sophisticated control of vocabulary and phrasing that is rarely imprecise; yet there are occasional weakness in idioms, colloquialisms, pronunciation, and cultural reference, or there may be occasional failure to interact in a totally native manner. (Has been coded S-4+ in some nonautomated applications.) (Data Code 46)

Speaking 5 (Functionally Native Proficiency)

Speaking proficiency is functionally equivalent to that of a highly articulate, well-educated native speaker and reflects the cultural standards of the country where the language is natively spoken. The individual uses the language with complete flexibility and intuition, so that speech on all levels is fully accepted by well-educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms, and pertinent cultural references. Pronunciation is typically consistent with that of well-educated native speakers of a some nonautomated applications.) (Data Code 50)

BEST COPY AVAILABLE

References

ACTFL. (1986). The ACTFL Proficiency Guidelines. Yonkers, NY: The American Council on the Teaching of Foreign Languages; and an earlier version, The ACTFL/ETS Provisional Proficiency Guidelines. (1982). Princeton, NJ: Educational Testing Service.

Buck, K. A. (1984). Nurturing the hothouse special: When proficiency becomes performance. *Unterrichtspraxis, 17(2).*

Buck, K. A. & Forsythe, G. (1985). Proficiency applications beyond the academic classroom. In C. J. James (Ed.), *Foreign language proficiency in the classroom and beyond* (pp. 147-164). Lincolnwood, IL: National Textbook.

Canale, M. & Byrnes, H. (Eds.). (1987). *Defining and developing proficiency: Guidelines, implementations, and concepts.* Lincolnwood, IL: National Textbook.

Child, J. R. (1987). Language proficiency levels and the typology of texts. In H. Byrnes & M. Canale (Eds.), *Defining and developing proficiency: Guidelines, implementations, and concepts* (pp. 97-106). Lincolnwood, IL: National Textbook.

Child, J. R. (1990). Textual norms across languages. Paper given to the ATSP in Miami Beach, FL, August 9, 1990.

Chomsky, N. (1975). *Reflections on language.* New York: Pantheon Books.

Clark, J. L. D. (1979). Direct vs semi-direct tests of speaking proficiency. In E. J. Brière & F. B. Hinofotis (Eds.), *Concepts in language testing: Some recent studies.* Washington, DC: TESOL.

Clifford, R. T. & Clark, J. L. D. (1987). The FSI/ILR/ACTFL Proficiency scales & testing techniques: Development, current status, and needed research. In A. Valdman (Ed.), *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency.* Bloomington, IN: Indiana University.

Defense Language Institute. (1992). Defense Language Proficiency Test IV: Familiarization Guide (DLIFLC Pamphlet 350-14). Monterey, CA: Author.

Defense Language Institute. (1982). *ILR handbook on oral interview testing.* Monterey, CA: Author.

de Jong, J.H.A.L. & van Ginkel, L. W. (1992). Dimensions in oral foreign language proficiency. In L. Verhoeven and J. H. A. L. de Jong (Eds.), *The construct of language proficiency* (pp. 187-205). Amsterdam: John Benjamins.

Douglas, D. (personal communication, June 15, 1993).

Herzog, M. & Shaw, V. (1983, December). *An indirect test of oral proficiency.* Paper presented at the meeting of the Modern Language Association.

Higgs, T. V. & Clifford, R. T. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher (p. 63)*. Lincolnwood, IL: National Textbook.

Lowe, Jr., P. (1986). Proficiency: Panacea, framework, process? A reply to Kramsch, Schultz, and particularly, Bachman and Savignon. *MLJ, 70*, 391-397.

Lowe, Jr., P. (1984). "The" question—*Foreign Language Annals, 17* (4), 387-388.

Authors

JAMES CHILD, RAY T. CLIFFORD, and PARDEE LOWE, JR. are interested in foreign language testing and evaluation.

457

Applied Language Learning
1993, Vol. 4, Nos. 1 & 2, pp. 55-63

Incidental Vocabulary Acquisition in French as a Foreign Language

Beatrice Dupuy
University of California, Los Angeles
and
Stephen D. Krashen
University of Southern California

Third semester college students of French in one intact class saw the first five scenes of Trois hommes et un couffin and read the next five scenes in class. They were then surprised with a vocabulary test that contained highly colloquial words that were in the text. Subjects performed significantly better than control subjects who were enrolled in another 3rd-semester French class as well as controls enrolled in a more advanced class, confirming that incidental vocabulary acquisition is possible in a foreign language situation. A conservative estimate of their rate of vocabulary acquisition was about .075 words per minute, which included the film and the reading. Rate of incidental vocabulary acquisition may have been underestimated, however, because the text was difficult, only 30 words were tested, and only eight of the 30 words appeared in the film.

A number of studies in first language acquisition have shown that students can acquire vocabulary through reading (for a review, see Krashen, 1993). In these studies, subjects were given a text to read that contained words unfamiliar to them. Subjects were not told to try to learn the new words, but were simply asked to read for meaning. After reading the text, subjects were surprised with a test on the new words. These studies have shown that readers make small but significant gains in vocabulary knowledge from even one exposure to a new word, gains that, given enough reading, appear to be large enough to account for adult vocabulary size (Nagy et al, 1985).

A small number of studies have been done showing that second language acquirers can also acquire vocabulary through reading (Pitts, White, & Krashen, 1989; Schouten-van Parreren, 1989; Day, Omura, & Hiramatsu, 1991). The purpose of the present study is to attempt to confirm that second language acquirers can acquire vocabulary incidentally and to get an estimate of how efficient this process is.

© 1993, Stephen D. Krashen & Beatrice Dupuy

Procedure

Subjects were 42 undergraduate students enrolled in French courses at the University of Southern California. Three intact classes were used: two intermediate French classes (one experimental and one control) and one advanced class (control).

The experimental class (n = 15), students in their third semester of French, saw the first five scenes of the film *Trois hommes et un couffin* (without subtitles). The next day they were asked to read the script of the next five scenes in class, approximately 15 typed pages of text. Since the text was considered to be challenging for students, the film was shown in order to provide some background knowledge. Subjects were given 40 minutes in class to do the reading. They were not told that they would be tested on vocabulary, nor were they told to pay special attention to new vocabulary. Instead, they were told they would be given a reading comprehension test based on the passage. Immediately after completing the reading, subjects took a vocabulary test (described below), but not the promised comprehension test.

The two control groups took the same vocabulary test, but did not see the film or read the passage. One control group (n = 13) was another 3rd-semester class, while a second control group (n = 9) was enrolled in a class in French pronunciation and conversation, which has 4th-semester French as a prerequisite. The second control group consisted of students who were more advanced than the experimental group and was included to confirm that the words on the test were unfamiliar to 3rd-semester students.

The vocabulary test (see appendix) contained 30 items from the reading that appeared between one and six times in the text. Words were selected from the reading by one of us (B.D.), a native speaker and instructor of French with extensive experience in teaching intermediate students. An attempt was made to select words that were extremely colloquial, and unlikely to be known by 3rd-semester students. A multiple-choice format was used, with all distractors in English. Each item had three possible answers as well as a fourth "I don't know" option, with correct answers randomly placed in first, second or third position. The reliability of the test was .956 (Kuder-Richardson formula 21).

We thus used a post-test-only design, without true randomization of subjects. This design is considered to be "pre-experimental" (Campbell & Stanley, 1966, refer to it as a "static-group comparison"). Its most obvious weakness is that subjects in the experimental and comparison groups might be quite different in their initial knowledge of the target vocabulary. This design, however, was our only option, since true randomization was impossible, and pre-testing might have cued the experimental group to focus on vocabulary acquisition.

There are, however, good reasons to presume that despite the use of this design, our results are valid. As noted above, nearly all the words tested were extremely colloquial and were probably not familiar to intermediate students of French (e.g. dingue, pôte, bosser); the second control group, consisting of more advanced students of French, was included to confirm that this was in fact the case. Also, as we will see below, the differences in test scores between the experimental and comparison groups were quite dramatic, which makes it unlikely that these differences were due to initial differences in vocabulary knowledge.

Results

As indicated in *Table 1*, the experimental group mean was clearly higher than either of the comparison group means. In addition, the performance of the comparison groups was much more variable, as indicated by the very different standard deviations in *Table 1*.

Table 1
Performance of Students who Read Trois hommes et un couffin *and Comparison Students on Test of Vocabulary*

	n	*mean*	*s.d.*
experimental	15	14.73	1.73
3rd-semester controls	13	8.08	5.24
advanced controls	9	8.93	5.71

perfect score = 30

Differences between the experimental and the two comparison groups were confirmed by planned comparisons, simultaneous *t* tests comparing the experimental group and each control group, with a Bonferroni correction for lack of independence. Because we performed a total of three *t* tests, a significance level of .016 was required for significance for individual tests (.05/3) (Pedhazur, 1982). Because the experimental hypotheses were directional, one-tailed *t* tests were performed.

For the experimental group and the 3rd-semester controls, $t = 4.04$, $p = .0003$; $df = 14$. For the experimental group and the advanced controls, $t = 3.66$, $p = .0012$, $df = 14$. (Because of the obvious difference in standard deviations of the experimental and control groups, a two-sample *t* was used, a version of the *t* test that does not assume equal standard deviations (Ryan, Joiner, & Ryan, 1985).

An interesting result was that the advanced controls performed similarly to the 3rd-semester controls, which confirmed that the target words were generally unknown to intermediate level students of French ($t = .40$, n.s., $df = 24$).[1]

Discussion

The experimental group clearly outperformed both control groups, confirming that vocabulary can be acquired incidentally by foreign language students.

A conservative interpretation of our results is that subjects gained an average of nearly six words after reading for 40 minutes and watching a film for 40 minutes (subtracting the experimental group score from the score of the highest-scoring comparison), or about .075 wpm, a little less than five words per hour.

These results, however, probably underestimate the value of comprehensible input in increasing foreign language vocabulary. Students found the text quite difficult; in fact, 30 percent of the experimental subjects did not finish reading the assigned passage in the allotted 40 minutes. An easier text might have resulted in even greater vocabulary acquisition. In addition, only 30 words from the reading were tested. It is quite possible that subjects acquired other words that were not tested. Finally, this estimate includes the time spent viewing the film, and most (22/30) of the target words did not appear in the film. Including only 40 minutes of reading would give an estimate of .15 wpm, quite close to the figure of .25 wpm calculated for rate of vocabulary acquisition for children reading in their first language (Nagy et al, 1985).[2]

This study did not attempt to demonstrate the superiority of vocabulary acquisition by reading, but it did confirm the results of previous studies showing that second language acquirers can acquire vocabulary incidentally. More careful studies, including the use of delayed testing, will help us determine just how efficient reading is. Even if it turns out to be the case that reading is not as efficient as direct vocabulary instruction, reading might still be preferable. Reading is far more pleasant than direct vocabulary instruction, and gives students many other benefits as well.

Notes

1. Two additional analyses were carried out, with similar results. In one analysis, we gave subjects credit for partial knowledge. For each test item, one distractor shared at least some semantic features with the correct answer. In the following example, a student who chose the distractor "annoyed" would receive partial credit (0.5):

renfrogné frowning
exhausted
annoyed
I don't know

When tests were rescored allowing for partial credit, the experimental group still outscored both control groups (experimental group versus 3rd-semester controls, $t = 3.99$, $p = .008$; $df = 14$; experimental group versus advanced controls, $t = 3.85$, $p = .0009$, $df = 14$).

We also reanalyzed the data, correcting for guessing, using the formula: *adjusted score = rights - (wrongs/number of choices - 1)*, assuming three choices (not counting "I don't know"). Once again, the differences between the experimental group and comparisons were significant (experimentals versus 3rd-semester controls, $t = 4.84$, $p = .0001$, $df = 15$; experimental versus advanced controls, $t = 3.50$, $p = .0015$, $df = 16$).

2. The rate of vocabulary acquisition estimated by Nagy et al was based on a scoring system that gave partial credit. Allowing partial credit (see *note 1*) raises our estimate of vocabulary acquired from reading from .15 wpm to .175 wpm.

Appendix

Vocabulary Test

1. douanier
   a) police dog
   b) customs officer
   c) cleaning lady
   d) don't know

2. en slip
   a) standing
   b) in briefs
   c) naked
   d) don't know

3. refrognée
   a) frowning
   b) exhausted
   c) annoyed
   d) don't know

4. moutard
   a) baby
   b) cyclist
   c) child
   d) don't know

5. enjoué
   a) happy
   b) down
   c) noisy
   d) don't know

6. dingue
   a) unfriendly
   b) funny
   c) crazy
   d) don't know

7. fouiller
   a) find
   b) search
   c) steal
   d) don't know

8. glousser
   a) yawn
   b) giggle
   c) chuckle
   d) don't know

9. biberon
   a) toy
   b) feeding bottle
   c) sucker
   d) don't know

10. bordel
a) dirty
b) pretty
c) messy
d) don't know

11. chiottes
a) closets
b) toilets
c) stairs
d) don't know

12. saccager
a) replace
b) smash
c) ransack
d) don't know

13. nana
a) mother
b) girl (colloq.)
c) baby
d) don't know

14. hargneux
a) surly
b) calmly
c) irascibly
d) don't know

15. crevé
a) tired
b) scrounged
c) excited
d) don't know

16. came
a) medicine
b) milk
c) drug
d) don't know

17. couffin
a) package
b) cradle
c) basket
d) don't know

18. couche
a) pillow
b) diaper
c) blanket
d) don't know

19. accablé
a) crushed
b) confident
c) numb
d) don't know

464

20. pôte
a) jug
b) buddy
c) boy
d) don't know

21. brique
a) broth
b) 1 million francs
c) drawings
d) don't know

22. engueuler
a) comfort
b) yell at
c) talk
d) don't know

23. taule
a) mole
b) cage
c) jail
d) don't know

24. abasourdi
a) astonished
b) worried
c) stunned
d) don't know

25. planouer
a) stuff
b) hide
c) throw
d) don't know

26. boulet
a) job
b) office
c) country
d) don't know

27. bosser
a) work
b) stay
c) toil
d) don't know

28. rot
a) burp
b) scream
c) toil
d) don't know

29. brailler
a) speak
b) squall
c) sleep
d) don't know

30. abruti
a) dumb
b) fat
c) gullible
d) don't know

References

Campbell, D., & Stanley, J. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental vocabulary learning and reading. *Reading in a Foreign Language*, 541-551.

Krashen, S. (1993). The power of reading. Englewood, CO: Libraries Unlimited.

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Pedhazur, E. (1982). Multiple regression in behavioral research. New York: Holt, Rinehart and Winston.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language, 5*, 271-275.

Ryan, B., Joiner, B., & Ryan, T. (1985). *MINITAB handbook*. State College, PA: Minitab.

Schouten-Van Parreren, C. (1989). Vocabulary learning through reading: Which conditions should be met when presenting words in texts? *AILA Review, 6*, 75-85.

Authors

BEATRICE DUPUY, Graduate student, Language, Learning and Literacy program, University of California, Los Angeles, CA 90089-0031. Specializations: vocabulary development, affective variables in second and foreign language acquisition

STEPHEN D. KRASHEN, Professor, School of Education, University of Southern California, Los Angeles, CA 90089-0031. Specializations: language acquisition theory, literacy development.

Applied Language Learning
1993, Vol. 4, Nos. 1 & 2, pp. 65-94

Instructional Implications of Gender Differences in Second/Foreign Language (L2) Learning Styles and Strategies

Rebecca L. Oxford
University of Alabama

This article addresses a potentially important topic in the language instruction field, gender differences in language learning styles and strategies. As background the author discusses some basic gender differences in human development and language use. The focus then shifts to research on styles and strategies in learning a second or foreign language (L2). Instructional implications for L2 classrooms are highlighted throughout.

Introduction

Do gender differences exist in styles and strategies for foreign and L2 learning, and if so, what instructional implications might these differences have? This is the key question. *Learning styles* are general approaches used in learning a new language or any other subject. Styles have cognitive, affective (emotional/attitudinal), physiological, and social aspects; they are not limited to the narrower concept of cognitive style (Keefe, 1979; Cornett, 1983). "Learning style [is] a consistent pattern of behavior but with a certain range of individual variability. . . . Styles, then, are overall patterns that give general direction to learning behavior" (Cornett, p. 9). Style dimensions include, among others, analytic and global; visual, auditory, tactile, and kinesthetic; and reflective and impulsive.

Learning strategies are more specific than learning styles. Strategies are the particular behaviors that learners employ, usually intentionally, to enhance their understanding, storage, retrieval, and ultimate use of information (Rigney, 1978). In the L2 area, examples include note-taking, seeking conversation partners, developing empathy with the target culture, skimming, previewing, and guessing to understand what is read or heard.

Specific themes of this article are:

1. *Gender differences in human development.* Information on social and cognitive development is necessary for understanding gender contrasts in L2 styles and strategies.

2. *Gender differences in language use.* Understanding how males and females converse in their own language (L1) and in a second language helps us comprehend gender differences in L2 styles and strategies.

3. *Gender differences in L2 learning styles.* This section synthesizes *numerous* studies, presents the author's own in-progress research on *gender* differences in styles in the L2 classroom, and offers instructional implications of style research.

4. *Gender differences in L2 learning strategies.* Existing research on L2 strategies is summarized here along with implications of this research for the classroom.

Gender Differences in Human Development

This section synthesizes research on gender differences in human social and cognitive development. The investigations mentioned here are representative of the main trends of research in this area.

Social Development

A review by Maccoby and Jacklin (1974) of hundreds of social development studies revealed that:

- Females show greater interest than males in social activities;
- Females prefer "gentle" interaction more than aggressive interaction (with aggression defined as the intent to harm others), and the opposite is true of males;
- Females are more cooperative and less competitive than males.

A meta-analysis by Hyde and Linn (1986) showed substantive gender differences in aggression (both verbal and physical) amounting to a .5 standard deviation, with males more aggressive than females. Even in early childhood, girls establish intense, nurturing, empathic relationships in pairs and triads, while boys travel in larger groups characterized by dominance-aggression hierarchies; these patterns continue through adulthood (Maccoby & Jacklin, 1974; Gilligan, 1982). Females are more likely than males to show a continuing need for approval and a desire to please others through good grades and social behavior (Bardwick, 1971; Tyler, 1965; Mansnerus, 1989; Nyikos, 1990; Docking & Thornton, 1979). They also smile and laugh significantly more than males (Hyde & Linn). See also Linn and Petersen (1985).

Cognitive Development

Developmental differences in verbal skills are sometimes pronounced (Good & Brophy, 1986; Halpern, 1986). In the native language, girls usually say their first words and learn to speak in sentences earlier than boys. This produces an initial "rate advantage" (Larsen-Freeman & Long, 1991). Later on, women often speak in longer, more complex sentences than men and score higher on tests of spelling, grammar, and perceptual speed. Boys have a far greater frequency of disabilities in learning, reading, and speaking than girls. On verbal ability tests and reading tests females on the average surpass males, particularly from age eleven on, according to most studies (Cahn, 1988; Gage & Berliner, 1975; Maccoby & Jacklin, 1974; Slavin, 1988). However, a few studies show a slight temporary drop in standardized verbal tests among girls during adolescence, possibly because of the recent inclusion of scientific and business topics (Nyikos, 1990; Mansnerus, 1989). See also Boocock (1990); Hyde (1981); Sherman (1978); Baker (1982); and Feingold (1992) for other views.

In learning an L2, males and females sometimes show different levels of skill, especially in listening. Eisenstein (1982) found that females surpassed males on dialect discrimination (listening) tasks and on recognizing greater and lesser prestige of dialects. Likewise, Farhady (1982), with a sample of 800 university students, discovered that females outperformed males in the area of listening comprehension.

Meta-analyses overwhelmingly indicate that, starting in adolescence, males generally outperform females in tests of mental object rotation with a difference of .26 to 1.0 standard deviation (Hyde & Linn, 1986). Males also typically do better on tests of depth perception, map reading, mechanical reasoning, and vertical-horizontal perception, and on mathematical tasks involving abstract concepts of space, relationships, and theory (Hyde & Linn, 1986; Maccoby & Jacklin, 1974; Weiner & Robinson, 1986), although not in ordinary arithmetic (Feingold, 1992). For other views, see Fennema and Sherman (1977); Caplan, MacPherson, and Tobin (1985); and Chipman (1988).

In reviewing existing research, Feingold (1992) found that scores for males were consistently more variable than for females in quantitative reasoning, spatial visualization, spelling, and general knowledge. According to Feingold (1992), trend analysis shows that gender differences among adolescents have decreased over the past generation.

Possible Causes of Gender Differences

Socialization is one of the main causes of gender differences in social tendencies and cognitive development (Rubin, Provenzane, & Luria, 1974; Smith & Lloyd, 1978; Block, 1973; Bem, 1974, 1975; Jacklin, 1983; Serbin & O'Leary, 1975; Nyikos, 1990; Eccles, 1989;

Slavin, 1988; Orlofsky & Stake, 1981; Mills, 1980; Crawford & Gentry, 1989). Socialization encompasses culturally defined gender roles and moral values, as well as exposure to particular courses of study and extracurricular activities.

In addition to socialization, physiology has been pinpointed as a source of many of these gender differences. Moir and Jessel (1991) argue that "brain sex" (anatomical differences in heterosexual male and female brains) causes differences in cognitive development. Gender differences in brain lateralization/hemisphericity have been noted, with greater nerve linkages between right and left hemispheres for females than males (Springer & Deutsch, 1989) and for homosexual males than heterosexual males (Elias, 1992; Associated Press, 1992). Probably any gender differences in development of social behavior and cognition arise from an intricate, not fully understood interaction of socialization and physiology.

What is the relationship between these gender differences in development on the one hand and learning styles and strategies on the other? Males and females, to a certain extent, have different tendencies both socially and cognitively. One hypothesis is that since we favor that which we are good at, these strengths become (general) learning style preferences, which are typically echoed in (specific) learning strategies.

Gender Differences in Language Use

Compared with men's native language speech, women's speech shows more empathy, concern, politeness, encouragement of other speakers, and more negotiation, detail-remembering, uncertainty, questioning behavior, and grammaticality (Fishman, 1978; Lakoff, 1975; Tannen, 1986, 1990; Kramarae, 1981). Men use more verbal expressions of power and aggression, adversarial-argumentative style, interruption, ridicule, analytical critique, and discouragement of other speakers (Dodd, 1991; Belenky, Clinchy, Goldberger, & Tarule, 1986, Tannen, 1986, 1990). Anthropologists Brown and Levinson and sociologists Giles and Tajfel attribute such L1 differences to socialization, with special reference to the subordinate role of women in economic and political spheres (see Kramarae, 1981, and Thorne, Kramarae, & Henley, 1983, for details).

Gass and Varonis (1986) studied the conversational behavior of learners of a second language. Men dominated conversations, but women initiated more negotiations of meaning, trying to understand and communicate clearly. These L2 patterns appear to reflect the general tendencies of men and women in native language use, as described above.

Gender Differences in Learning Styles

Research outside the L2 area has cited many gender differences in learning styles, and to some degree L2 researchers have started focusing on such contrasts. Let us take each relevant style dimension in turn, although some researchers (Schmeck, 1983; Willing, 1988) have postulated that all these dimensions might eventually yield a single global versus analytic contrast capturing a large majority of human variation in learning style.

Brain Hemisphericity

Brain hemisphericity or lateralization (right, left, and integrated) is a feature of many learning style surveys and books. Right-hemisphere dominant individuals—those whose right side of the brain typically dominates their thinking processes—tend to be more field dependent (less able to separate the details from a confusing background), global, and emotion-oriented, according to Willing (1988), while left-hemisphere dominant people are more field independent, analytic, and logic-oriented

Both hemispheres deal with language, but they do it differently. The left hemisphere interprets the meaning of words, while the right hemisphere interprets verbal tones, patterns, and musical qualities of the language.

Springer and Deutsch (1989) assert that in men, compared to women, the left hemisphere is more lateralized (specialized) for verbal activity and the right hemisphere is more lateralized for abstract or spatial processing; women use both the left and the right hemispheres for verbal and spatial activity. In women as compared to men, part of the corpus callosum (the bundle of brain fibers linking the left and right hemispheres) is bigger in relation to overall brain weight, allowing more information to be exchanged between the two hemispheres, thus lending strength to the possibility that women have more integrated brain function than men. See also Associated Press (1992) and Elias (1992).

Other research contradicts the idea that brain hemispheres are more integrated in females than in males. For example, some research implies that females are usually more right-hemisphere dominant than males because of their frequent field-dependent attributes, discussed later (Good & Brophy, 1986; Shipman & Shipman, 1985; Sigel, Jarman, & Hanesian, 1967). See also Hauck (1985).

Relevance of hemsiphericity to L2 learning:

Based on the research just noted, males might usually process the L2 more readily through the left-hemispheric, analytic mode, but females might more often process the L2 through the right-hemispheric, global mode or through an integration of left- and right-hemispheric modes. The implication is that mixed-gender L2 instruction should use aspects of both modes, thus including a variety of analytic activities and global-fluency tasks.

Leaver (1986) found that hemispherically integrated L2 learners performed well at all proficiency levels. Students who processed the L2 mostly through the right hemisphere were better learners at lower proficiency levels, where the focus is on intonation and rhythms, while those whose L2 processing was chiefly via the logical left hemisphere were better learners at higher proficiency levels, which call for greater control and analysis. Females in Leaver's sample, unlike typical females in our culture, were mostly left-hemisphere learners rather than integrated or right-hemisphere learners.

Field Independence and Dependence

Gender differences have often been found for field independence and dependence through many different measures: tests of perception in titled rooms, embedded figures tests, and rod-and-frame tests. As adolescents and adults, and sometimes even as children, males are usually more field independent than females (Good & Brophy, 1986; Shipman & Shipman, 1985; Dyk & Witkin, 1965; Witkin & Berry, 1975; see Lusk & Wright, 1981 for another view).

Field independence refers to the ability to separate easily the key details from an ambiguous context through the use of analysis. Research with all ages suggests that field independent people are less sensitive to the social context, are perceived as more detached and more logical than field dependent people, and prefer more structured, analytical learning (see review in Oxford, 1990a). Field independence is often linked with left hemisphere dominance.

Field dependence (more appropriately called field sensitivity) is the inability to separate details from the background easily and implies a tendency toward global impressions. Research shows that field dependent people tend to be more sensitive to the social context, are perceived as more outgoing and more considerate than their field independent peers, and do well with less structure in their learning

(Oxford, 1990a). Field dependence is seen to be related to the dominance of the right side of the brain.

Relevance of field independence/dependence to L2 learning:

The relationship between the field independence-dependence dimension and L2 is not clear-cut. Research has shown some advantage for field independent learners, often males, in L2 achievement (Hansen & Stansfield, 1981). However, this might be related to the analytic nature of most written L2 achievement tests and many course activities (as noted by Leaver, personal communication, November 24, 1991). Field dependent individuals, often females, with their more interpersonal and global orientation, might have an edge in less analytic aspects of overall L2 communicative competence, such as sociolinguistic competence, discourse competence, and strategic competence. However, more evidence is needed to substantiate the possible link between field dependence and these aspects of communicative competence (Galloway & Labarca, 1991).

If the students' degree of field independence is not known in a mixed-gender class, the best option is to provide a combination of both more structured, analytical activities and less structured, globally communicative activities.

Ways of Knowing

Based on a detailed, interview-based study of college students, Belenky, Clinchy, Goldberger, and Tarule (1986) argue that in general, men and women have different ways of knowing, men more often through "objectivity and thinking" (abstract analysis) and women more often through "subjectivity and feeling" (personalized experience).

This assertion is aligned with years of research using the Myers-Briggs Type Indicator (MBTI). The MBTI Data Bank, covering hundreds of studies, reports that 60 percent of women in our culture show more a feeling (values-and-emotions-based) approach than a thinking (analysis-founded) approach (Myers & McCaulley, 1990). The opposite proportions appear for men: 60 percent of men in our culture demonstrate more of a thinking approach than a feeling approach. By definition, the feeling-related style is more personal, and the thinking-related style is more impersonal. People who have a *feeling style* show greater concern for other people's feelings and values, while those with a *thinking style* are more concerned with objective standards of rules, facts, and justice (Myers & McCaulley, 1990; Lawrence, 1984).

Relevance of "ways of knowing" to L2 learning:

Extrapolating from the research above, we can say that males and females, in general, might go by slightly different routes when learning a new language. In L2 learning, more males than females might favor the thinking approach (rules, facts, logic, objectivity, impersonal or less interactive learning), and more females than males might prefer the feeling approach (feelings, cultural sensitivity, empathy, subjectivity, socially interactive or cooperative learning). More research is needed here.

Reflection and Impulsivity

The women in the Belenky et al. study wanted to reflect before judging; the men in the study, on the other hand, rapidly jumped to conclusions. "Even when they [the women] disagreed with an opinion, they hesitated to judge it wrong until they had tried to understand the reasoning behind it" (p. 116); most of the men did not delay so as to understand better. The nonjudgmental stance shown by the women in the study might reflect some learning-style attributes of reflection (as contrasted with impulsivity). The nonjudgmental stance might also reflect the politeness and social empathy characteristic of most women. Note that in conversational language use, as mentioned earlier, men typically interrupt and dominate, while women ask more questions and thus elicit more information from the other person; these tendencies might be echoed in the impulsivity versus the reflection research findings.

In psychological research, reflection is defined as the tendency to stop and consider options before responding, often resulting in greater accuracy, while impulsivity is the tendency to respond immediately and often inaccurately. (Note that fast-accurate and slow-inaccurate are not part of the typical research paradigm of reflection and impulsivity.) In one study, gender differences appeared significantly in reflection-impulsivity for seven-year-olds, with females being more reflective; impulsivity was found to be related to poor reading skills (Roberts, 1979). In studies with children and adolescents, Shipman and Shipman (1985) found that whenever gender differences appeared, females were usually more reflective.

Relevance of reflection/impulsivity to L2 learning:

In L2 research by Ehrman and Oxford (1988, 1989, 1990), the perceivers on the MBTI (those learners who could refrain from leaping to conclusions too rapidly and who might therefore be

called reflective learners) were better at learning languages than the judgers *(those who needed quick closure). The perceivers/reflectives stayed open to gather new clues which would help them to understand the meaning. Likewise, in other L2 research impulsive learners had problems because of their premature, inaccurate responses; an imposed wait time was helpful (Parry, 1984; Galloway & Labarca, 1991; Birckbichler, 1984).*

However, too much concern for accuracy can become destructive anxiety, diminishing L2 performance, as Ehrman and Oxford (1990) found empirically. If carried too far, reflection in an L2 situation can be immobilizing, especially in a fast-moving, intensive program (Leaver, personal communication, November 24, 1991). A balance between reflection and impulsivity appears desirable for L2 learning, especially when conversational fluency—not just accuracy—is an important focus.

Some nagging questions remain about gender differences in reflection and impulsivity. Females are often seen as more reflective than males (see the Belenky et al. study and the Shipman & Shipman review.) If this is true, what are we to make of the frequent assumption that reflection is associated with analytic, field independent people, who are, according to the research, more often males than females? And how does this fit in with the finding that females compared with males are often more accurate in spelling and grammar—skills that, at first glance, might seem thoroughly analytic?

Perhaps there are really two kinds of reflection operating in the L2 field. One type might be analytic reflection that allows the person to apply grammar and spelling rules logically, to carefully analyze sociolinguistic factors in order to produce the appropriate response and to conduct detailed self-evaluation. Another type might be global reflection that enables the learner holistically and correctly to use patterns of grammar and spelling, instinctively comprehend social interactions and situations without analysis, and conduct global self-evaluations. Both kinds of reflection are potentially important. It is entirely possible that these different types of reflection, if well researched, would show gender differences that might help

BEST COPY AVAILABLE

answer many remaining questions about reflection and impulsivity.

Sensory Preference

Many style studies focus on sensory preferences, sometimes known as perceptual modes or learning channels: visual, auditory, tactile, and kinesthetic. *Visual* students are those who prefer to learn via the visual channel: therefore they like to read, need the visual stimulation of bulletin boards and movies, must have written directions, and often prefer working alone. *Auditory* students are those who enjoy the oral-aural learning channel; thus, they want to engage in discussions, conversations, and group work and typically require only oral directions. *Tactile* students are those who need to touch and manipulate objects; and *kinesthetic* students are those who require movement and frequent breaks in activity.

One might speculate that a learning style preference characterized by tactile and kinesthetic channels might be related to the spatial ability that is prominent in the masculine gender role. This indeed seems to be the case in a study reported by Hansen (1982). There is little information on gender differences in visual or auditory preferences.

Relevance of sensory preferences to L2 learning:

In L2 classrooms we might expect that tactile and kinesthetic students might more often be males than females. We might also predict that nontraditional females (women who have chosen to follow gender-role patterns different from the traditional norm) would show these preferences more frequently than traditional females. L2 activities, for both children and adult learners, should involve some movement for kinesthetic people and manipulation of objects for tactile people.

Linkages between gender and auditory preferences are not clear, although two listening studies (Eisenstein, 1982; Farhady, 1982) suggest that L2 auditory ability might be greater in females than in males. Students who are less auditory, often males, can be taught to develop listening strategies (see Mendelsohn, 1984). The relationship between gender and visual preference in the L2 setting has not yet been determined.

Because L2 learning requires an integration of the senses, especially visual and auditory, L2 teachers must help this integration to occur. Making this link might involve using multiple media—print, pictures, objects, sounds—all related to the same theme or task.

Findings from a Study in Progress: Examples of Cross-gender Learning Style Conflicts

Regarding gender differences in learning styles, here are some early results of an in-progress investigation by this author, along with Roberta Lavine of the University of Maryland and Sharon Fechter of Mount Vernon College. The 150 students involved in the study were from graduate and undergraduate classes at two major state universities and a small four-year liberal arts institution. Participants were asked to describe a style conflict they experienced with a language teacher (L1 or L2) during their learning experiences in secondary school or university. If no style conflicts had occurred, participants were to so indicate. Only about two percent of the participants said they had experienced no such style disparities with their teachers. Before writing their narratives, each of the students had either taken at least one learning style survey (listed by Schmeck, 1983) or participated in a class discussion to identify their own learning styles. Although the students had not become style experts, they identified their own styles in a way that was valid to their teachers, who were specialists in learning style research. The students' "style war" narratives reflected a sound understanding of basic style concepts and the psychological labels attached to these concepts.

Eight examples were chosen for discussion here. Examples 1, 2, and 3 were from advanced graduate students in L2 education, and all three had significant teaching experience themselves. Examples 4 and 5 were from undergraduate Spanish language students. Examples 6 and 7 were from undergraduates in English education. Example 8 was from a graduate student in English education. The headings for each example summarize the salient style dimensions found directly in the participant narratives. Note that in these examples, the males and females do not necessarily exhibit styles typical of their gender; instead, the examples were selected to illustrate a range of different style conflicts. More detailed content analyses, along with relationships between narrative style results and student grades, will be reported at a later date.

The verbatim narratives describe learning style conflicts between students and teachers in language classes. All these conflicts involved a difference in gender between the student and the instructor. Most instances of conflicts did not specifically cite gender as a major issue. However, when style battles occurred between a teacher and a student of different genders, the frustration level and inability to communicate appeared to be higher than when the style conflicts were single-gender. Thus, a style conflict which might or might not have been directly related to gender was often exacerbated by a gender-related communication block.

Example 1. Female global/impulsive/visual graduate student and male analytic/reflective/nonvisual teacher:

My style wars began in a graduate course. The teacher and I are not compatible at all. My style of language learning is causing great friction with his teaching mode. I'm unhappy, I dread the class, and I'm under a great deal of stress in and out of the class.

I'm a visual learner . . . The teacher never writes anything on the board; in fact, he never moves from his domain, the chair. I'm always prepared for the day's work, [but] I like to have fun while I'm learning. My teacher is close-minded. Everything has to be done his way. He is dogmatic . . . He has taught this way for fifty years and will continue to do so. I'm a global learner; he likes to have everything detailed at length with no explanation given. I'm having to use rote memorization in order to study, and that is no joy; nothing has meaning, everything is just a blur. He is slow, and I'm impulsive, but not to the point I do not take my studies seriously. He uses overcorrection methods which really dampen the atmosphere of learning in the class . . . I do hope that I will make it to the end of the course without an ulcer or have a nervous breakdown before the exam. This has really been a nightmare and a very forceful style war in my learning career.

Example 2. Female analytic/closure-oriented graduate student and male global/open teacher:

I took a Portuguese class in graduate school in which I experienced "style wars." My professor was an intuitive-perceiver [open] and global teacher. There were many open-ended activities and the class was somewhat chaotic; but fun, I must admit.

I am an intuitive and closure-oriented student. I like organized presentations and detailed grammar explanations. This professor was very animated in the classroom, and he used drama and role-plays to teach grammar inductively. We watched a videotape that accompanied the textbook and we learned songs and cultural points of interest.

We did a lot of oral/aural work in the language laboratory. When I finished the course, I wondered what I really learned. I could not reproduce any verb conjugations or many vocabulary words from memory. I enjoyed the class a great deal, but was uncomfortable with his teaching style. I did not learn in an environment with which I was comfortable.

Example 3. Female introverted/visual Korean graduate student and male extroverted/haptic (tactile/kinesthetic) American teacher:

I am an introvert, sensing, feeling, and judging (closure-oriented) learner. I am also strongly visual. The reason I have [a] strong visual learning preference is mainly due to my previous learning experiences in Korea, where most teachers emphasize learning through reading and tend to pour a great deal of information onto the chalkboard. I once had a style conflict with one of my professors at a California university, where I took an elementary communication class. The professor was a strong haptic (tactile/kinesthetic) person who always wanted his students to demonstrate certain kinds of motions and actions in the classroom. The professor himself liked to act, dance, mime, and move around the classroom. He rarely wrote down his lecture points on the chalkboard . . . Since I am a very strong visual learner, his lecture style made me feel very uncomfortable and insecure . . . The worst part happened a few weeks after the beginning of the semester, when every student was grouped and required to work in groups outside the classroom . . . The activity book we were to use was full of interpersonal communicative activities requiring "too much" body movement at the same time.

I have no doubt that many of the activities that the professor developed were excellent for the course. Personally, however, I had difficulties adapting to such activities, firstly because I was highly introverted so that I didn't feel comfortable working in groups, and secondly because I was not a haptic learner . . . I often skipped the required group activities. Frankly, I did not enjoy the course at all.

Example 4. Male reflective/introverted undergraduate student and female impulsive/extroverted teacher:

This class [is a style war]! Impulsive teacher, reflective me. That's why this class isn't much of a learning experience for me. When we review the homework, it's "Let's do this quickly, very, very quickly." It's not like we're going to learn anything more if we move at a faster pace. I also don't like working in groups very much. Not because I don't want to work with other people, but if what we're working on is something new (which it ought to be), then its like the blind leading the blind. On top of this, [it] is hard to work with other people who "work" differently.

Example 5. Female visual/extroverted undergraduate student and male nonvisual/introverted teacher:

One example of when I truly did not *find it easy to* learn *in* Spanish class*—and it was quite a shock—was that during eighth grade, my teacher used* no *visuals, and had* very *little class interactions whatsoever. It was certainly* not *the people, because everyone was very social and interactive ordinarily. He just made* no *effort to enliven the class. The only thing that he would do is to stimulate us with a* Spanish film*: but this would fail miserably because we would not respond* back *to a film, or ask* it *questions. We wound up with a* stagnant *group of students; and I'm sure nobody* learned *very much. [Emphasis in original.]*

Example 6. Male tactile/kinesthetic undergraduate student and female teacher who was not tactile/kinesthetic:

The battle I fought . . . dealt with my tenth grade teacher. Her style of rigid lecture conflicted with my learning style of hands-on [tactile/kinesthetic] and investigative learning. This caused some problems between the two of us. My attention span in her class was very low, and her tolerance span of me was about as low. Consequently, I only did average work in her class, while in others I did much better.

Example 7. Female analytic/closure-oriented undergraduate student and male global/open teacher:

My conflict was in a 17th century literature class. I'm very analytic and really like to know when things are due and what is expected of me. He [the teacher] did not give us a syllabus until after the first two class meetings. Then he rambled in his lectures; there seemed to be no continuity.

Example 8. Male global graduate student and female analytic teacher:

The only style conflict I have ever had in an English class was in the tenth grade. My teacher was very analytical, whereas I am very global. She took every aspect of the class step by step. She never deviated from her plan. She refused to see more than one interpretation for a situation. We had an extremely difficult time communicating, thus I had a hard time in learning because she could not adapt to another style.

Many of the other style-conflict narratives in the sample of 150 were gender-related. It was striking that males who described style conflicts often spoke in somewhat impersonal, objective terms. Females, on the other hand, focused closely on particular people, told exactly what those people did, and gave a strong emotional reaction. This suggests a somewhat stronger focus among women than men on very personal and emotional aspects of learning and teaching.

Relevance of style conflicts to L2 learning:

Style wars exist more often than L2 teachers might like to believe. Sometimes we attribute them to laziness, stupidity, or lack of motivation on the part of the L2 student, but more often the situation is far more subtle. Style conflicts in L2 classes do not always involve a difference in gender between student and teacher, but when they do, the conflict might be more difficult. Suggestions on how to identify and cope with style issues, including style conflicts (gender-related or otherwise), are given next.

General Implications of Style Results for the L2 Classroom

1. *Assessing Styles.* It is important to assess the learning styles of L2 students and acquaint these students with their preferred styles. Teachers should also assess their own styles. This can be done using any of the wide array of learning style instruments (see Schmeck, 1983). Using more than one of these instruments to measure style often gives a more complete picture. Sharing information about what the instruments measure can demonstrate the teacher's interest not only in what students learn but also in how they learn. This can be part of general *learner training*, which has been shown to be of significant benefit to L2 learners (Wenden, 1991; Ellis & Sinclair, 1989; Brown, 1989, 1991; Cohen, 1990; O'Malley & Chamot, 1990; Oxford, 1990b).

2. *Accepting Gender-Related Differences.* Teachers and students should not be surprised if gender-related differences appear in the style assessments. All participants should feel free to discuss these differences openly, bringing up any sociocultural (and perhaps biological) influences that might have helped create style contrasts between females and males. It is useful to emphasize the number of useful skills people of different styles can learn from each other, both across and within gender boundaries. In addition, cross-gender likenesses in style should also be highlighted in the discussion.

3. *Using Style Information for Tailoring Instruction.* By using style data, teachers are better able to spot style conflicts in the classroom and

BEST COPY AVAILABLE

help students move beyond their basic style tendencies when necessary. Teachers can also vary their instructional techniques to meet the needs of students with contrasting styles of learning. Providing a wide range of activities—analytic and global, thinking- and feeling-oriented, visual and auditory—and appealing to all possible styles is the best preventive medicine for L2 style problems. Teachers can try the "imposed wait time" of 30 seconds to require or encourage impulsive or conversation-dominating students, often males, to reflect before responding, but they can also provide time-dependent games and skits that make good use of these students' rapid answering skills. Teachers can give reflective students, often females, ample opportunity to think and ask questions before responding, but they can at other times consciously encourage these students to be more spontaneous through multiple activities in small groups. Teachers can offer myriad multisensory activities that allow tactile/kinesthetic students, who may preponderantly be males, to move around and handle objects while learning the language; and they can offer visual and auditory students the different kinds of sensory stimulation they need. No matter what their gender, students are themselves likely to find information about their personal styles valuable in L2 learning, because this information helps them seek activities and learning environments suitable to their particular learning needs (Oxford & Lavine, 1991).

4. *Preparing the Learning Environment.* Style data can help the teacher prepare a learning environment that welcomes and accommodates a variety of styles for both females and males. The learning environment can establish the class as either exclusive, limiting acceptable styles, or inclusive, welcoming many styles. Lively bulletin boards, attractive exhibits of cultural artifacts (especially touchable or manipulable items), eye-catching wall decorations, videos, tapes, creative use of space with areas for movement activities, comfortable seating (a small sofa or cushions along with regular desks) and flexible seating for private reading or group activities can send positive, tacit messages to students: "Learning is enjoyable here," "Males and females are equally welcome," and "There are different ways to learn in this classroom."

So far this article has explored gender differences in development, in language use, and in learning styles with applications to L2 settings. Information is also available about gender differences in learning strategies, discussed next.

Gender Differences in Learning Strategies

Learning strategies, the specific behaviors students use to improve their learning, are important to achievement in any subject area (Oxford, 1989). Non-L2 researchers have discovered that effective learners actively associate new information with existing information in long-term memory, and that the use of well-chosen strategies separates experts from novices in a variety of fields: physics, native language reading, and engineering, to name a few.

In the L2 area, almost two decades ago researchers made lists of strategies presumed to be essential for the "good language learner." Rubin (1975) suggested that good language learners (1) guess willingly and accurately, (2) want to communicate, (3) are uninhibited about mistakes, (4) focus on both structure and meaning, (5) take advantage of all practice opportunities, and (6) monitor their own speech and that of others. Another early list (Naiman, Frohlich, & Todesco, 1975) added that successful L2 learners think in the language and use affective (emotional) strategies to cope with the strains of language learning. More recent L2 researchers found that appropriate strategy use is often related to better learning performance (Ehrman & Oxford, 1988; Wenden & Rubin, 1987; Wenden, 1991; O'Malley & Chamot, 1990; Cohen, 1990; Brown, 1989, 1991). Mere frequency of use of learning strategies is not as important as orchestrating the use of strategies to fit the requirements of the task and the goals of the learner. Unsuccessful L2 learners often use as many strategies as effective L2 learners, but in a random, almost desperate way, as suggested by Vann and Abraham's (1989) research. Strategy training studies have shown that L2 performance can frequently be improved through teaching students to use new strategies or to improve their application of familiar strategies (O'Malley & Chamot, 1990; Oxford & Crookall, 1989).

Just a few studies (e.g., Cox & Waters, 1986; Wandersee, 1988) outside the L2 field have investigated gender differences in learning strategies, and these have produced conflicting results. Surprisingly, the most comprehensive volume on non-L2 learning strategies (Weinstein, Goetz, & Alexander, 1988) entirely ignores the topic of research on gender differences in strategy use.

Only recently have researchers noted that the choice of L2 learning strategies is often gender-related and depends crucially on L2 learning style, which is also frequently linked to gender (Ehrman & Oxford, 1988, 1989; Oxford & Ehrman, 1988; Oxford, 1990a, Oxford 1990c, Nyikos, 1990; Willing, 1988). If learners are not pressed by the situation or the teacher, they tend to employ L2 learning strategies that are directly congruent with their favored learning style. For instance, the analytic, field-independent learner (probably a male) ordinarily selects strategies involving logic, such as deductive reasoning, whereas the global, field-dependent learner (often a female) chooses nonanalytic

strategies that involve searching for the main idea and intuitively guessing from multiple contextual clues—frequently social ones—when some pieces of information are missing. The visual learner (male or female) seeks extensive visual input through reading; the auditory learner (frequently a female) engages in lots of social interaction and uses auditory memory devices like rhyming; and the tactile/kinesthetic learner (typically a male) likes total physical response techniques and manipulating objects. The reflective learner (often a female) considers different angles as well as the social context before responding and is devoted to answering correctly, while the impulsive learner (often a male) jumps in with a quick response and may want to dominate, regardless of the correctness of the response. These somewhat gender-stereotypic descriptions do not imply that all females or all males prefer strategies; these illustrations do, however, show that the favored style dimensions, which are frequently gender-related, can have a direct effect on the specific L2 learning behaviors chosen.

Oxford, Nyikos, and Ehrman (1988) were the first to publish a review of studies involving gender differences in L2 learning strategies. At that time, out of a selection of over 80 investigations of L2 learning strategies, only four made any mention of gender differences. Since that time, more research (discussed later) has been done on this topic.

The first of the four studies described in the 1988 review was by Politzer (1983), who reported that female college students used social strategies for L2 learning significantly more often than their male peers—a difference which Politzer did not explain but which might be related to gender differences in social orientation (see the discussion above on development). In the second study reviewed, Ehrman and Oxford (1988) found that females in an intensive adult learning setting, compared with males, reported significantly greater use of L2 learning strategies in four categories: general study strategies, strategies for negotiating meaning, self-management strategies, and functional practice strategies. The third study discussed is that of Oxford and Nyikos (1989), who found that female college students, in contrast to males, used L2 learning strategies significantly more often in three of five categories derived factor-analytically: formal rule-based strategies (see Note 1), general study strategies, and conversational input-elicitation strategies. The gender differences found in the second and third studies might be associated with women's social skills, stronger verbal skills (including pattern usage), and greater conformity to academic and linguistic norms. The fourth study reviewed was by Nyikos (1987), who discovered significant gender differences among university students in using memory strategies for German vocabulary learning. After training in the use of these strategies, men outperformed women in the color-plus-picture combination, which was explained as potentially related to men's putatively greater visual-

spatial acuity. However, women surpassed men in the color-only condition, which was explained by women's greater interest in color (often seen as a social attractor).

Following this 1988 review, other studies focused on gender differences in L2 learning strategies. Oxford, Park-Oh, Ito, and Sumrall (in press) discovered gender-difference trends among high school students studying Japanese by satellite. Although all students used various strategies, females tended to use cognitive strategies, social strategies, and affective strategies more often than males; gender differences were not as strong in metacognitive, compensation, and memory strategies. Males in this study, as in others, did not surpass females in strategy use in any of the main strategy categories. Females also outperformed males in terms of motivation and Japanese language achievement.

In sharp contrast with these findings, Tran (1988) studied gender differences in English language acculturation and learning strategies among Vietnamese adults over age 40 in the United States. He found that females had more L2 learning problems and that males were more likely than females to use a variety of L2 learning strategies to improve their English skills. Possibly age and cultural gender-role differences influenced these results.

In a study of 500 ESL learners in Australia, Willing (1988) discovered interesting gender differences in L2 learning strategies. Significant differences included the following: women more often than men indicated they liked to learn many new words (no sensory modality implied); to learn English words by seeing them (visual modality implied); to learn English words by "doing something" (kinesthetic); and to learn by talking with friends in English (auditory). Men more often than women said they wanted to write everything in their notebooks (visual and kinesthetic) and liked to learn by using cassettes (auditory), although these two strategies were rather unpopular among both men and women. Thus, more of the significant differences in strategy choice, when they occurred, favored females, but these differences did not clearly separate females from males on the basis of any underlying sensory preferences.

Relevance to L2 learning:

Whenever L2 strategy research has considered gender, it has usually demonstrated gender differences in strategy frequency, with females choosing to use particular sets of strategies more often than males. Females especially tended to use general study strategies, social strategies, affective strategies, and certain conversational or functional practice strategies more frequently than males across a number of studies, thus usually showing a greater range of frequently used strategy categories. Males and females favored contrasting memory strategies in a study of vocabulary learning. Differences in frequency and patterns of strategy choice might explain why females often show better L2 classroom performance than males (although of

course informal, non-classroom based language development in the target country might reveal something different if research were conducted).

General Implications of Strategy Results for the L2 Classroom

1. *Assessing Strategy Use:* Teachers should measure the learning strategy use of L2 students and help students become aware of the strategies they habitually employ. Oxford (1990b) lists a number of useful strategy assessment techniques: surveys, think-aloud procedures, diaries, interviews, and observations. With some exposure to strategy assessment, students typically become intrigued about their own L2 learning behaviors, and they like to know what strategies are typical of their gender and what strategies are used equally by females and males. Moreover, students usually want to develop new strategies or improve their use of existing ones.

2. *Conducting Learner Training:* Style-and-strategy training can help all students, male and female, understand their styles and strategies and can enable students to make better decisions about how to meet their own learning needs and fulfill the requirements of any L2 task. This training can be provided by the teacher in the context of the ongoing L2 class. It can be woven into regular class activities in a natural but explicit way. For instance, the teacher might provide a range of memory strategies (such as imagery, rhyming, total physical response, repetition, and context-creation) for learning vocabulary words when they appear in the lesson and might explicitly point out that these are useful strategies. The teacher might also indicate that students might prefer particular types of memory strategies, e.g., visual students would prefer imagery, auditory students would prefer rhyming and choral repetition, and tactile students would like labeling real objects.

3. *Stretching:* This training can also stretch students beyond the strategies that fit their natural style preferences. For example, teachers can help global, feeling-type students (those who make decisions based on emotions and personal values)—often females—to make the most of their empathetic tendencies in conversational settings and can also provide them with new analytic strategies (e.g., word dissection, deductive reasoning) to figure out meanings logically. Teachers can provide plenty of avenues for analytic students—frequently though not always males—to ply their logical trades, but can also stretch these students in the direction of greater holistic, sociocultural sensitivity and synthesizing ability.

Summary and Conclusions

This article has described some key gender differences in styles and strategies, with primary focus on the learning of a second or foreign language. Anatomy is not destiny, contrary to Freud's dictum, but gender appears to have some profound effects on the ways that males and females approach L2 learning. In this paper, gender-related differences in development and the use of language offer a backdrop to the presentation of gender differences in L2 learning styles and strategies. Women and men often show somewhat different patterns of social and cognitive development, and their uses of language in conversation appear to contrast rather strongly.

In sum, some existing research on learning style suggests that males might have more strongly specialized brain hemispheres than women, the left hemisphere in men being dominant for language. Some research suggests that, compared with men, women more readily use the right hemisphere or use both hemispheres in an integrated way for language processing. In processing language, men are often somewhat more field independent, analytic, objective, and logical-minded, while women have been found to tend toward field-dependent, globally patterned, subjective, and emotional capabilities. In many investigations females have been more reflective than males. However, the distinct possibility of multiple kinds of reflection has yet to fully surface in psychological and educational research, particularly in the area of L2 learning. Some auditory style preferences have been shown by females more than by males. Preliminary research shows that style conflicts between teachers and students occurred frequently and were exacerbated by faulty cross-gender communication.

L2 learning strategy research shows that females reported using more varied strategy types and employing strategies more frequently than males. In the available studies, females more often than males reported tendencies toward using general study strategies, social and emotional strategies, and conversational or functional strategies. Across most studies, males failed to report greater strategy frequencies in any of the major strategy categories.

What are researchers' responses to gender differences like these? Six kinds of recommendations have emerged recently in psychology and education journals:

1. *Ignore gender differences for political reasons.* Some researchers, such as Baumeister (1988), suggest that gender differences should not be reported at all, because such reporting might perpetuate and legitimize the gender distinction, which then might justify continuing the unequal treatment of women and men by society. This seems to be a head-in-the-sand position. The presence of gender differences does not at all justify

social injustice or inequality of opportunity, and articles about gender differences should emphasize this fact if there is any question about it.

2. *Ignore gender differences because some might be small.* Chipman (1988) calls for researchers to ignore gender differences on the grounds that many of these differences, though statistically significant, are small in magnitude. Yet Chipman herself cites meta-analyses in which some gender differences (e.g., aggression and spatial rotation ability) are relatively large. Thus, her argument seems less than fully convincing.

3. *Emphasize similarities between males and females.* Rothblum (1988) suggests that nonsignificant and small gender differences should be published to emphasize similarities between males and females. If this argument is taken to mean that similarities should be highlighted as well as differences, the argument appears reasonable. If only similarities and no differences were reported, then the suggestion would appear politically motivated and not useful.

4. *Examine gender differences in terms of variability, not just in terms of average performance.* Feingold (1992) urges researchers to look at gender differences not only in means but also in variability. He points out, as we have noted, that males show greater variability than females in a number of cognitive areas. Feingold's suggestions seem entirely appropriate and deserve prompt attention by researchers.

5. *Within traditional quantitative research, describe gender composition, generalize appropriately, and replicate studies.* Researchers should specify the number of males and females in every study, according to guidelines of the American Psychological Association (as noted by Gannon, Luchetta, Rhodes, Pardie, & Segrist, 1992). Lack of such specification has been one reason for inappropriate generalizations of findings in one-third of the 5,000 articles published in prestigious psychology journals in the last twenty years (Gannon et al.). Because many L2 learning articles fail to describe the sample adequately in terms of gender and other demographic features, such inappropriate generalizations can easily occur in the L2 field as well. Improved adherence to the ordinary tenets of traditional empirical research would allow reasonable generalizations to be made and, just as importantly, would allow replication of gender-related research findings across studies (Denmark, Russo, Frieze, & Sechzer, 1988).

6. *Create a better balance of qualitative and quantitative research related to gender.* Riger (1992) and many others call for science to honor subjective, naturalistic, socially contextualized, qualitative research on an equal footing with objective, experimental, socially decontextualized, quantitative research. Some people call this dual-aspected science a

"new feminist epistemology" (Riger, 1992), but it is in effect a more realistic and comprehensive human epistemology. More efforts in qualitative and naturalistic research would provide a broader context for understanding quantitative gender differences when they occur. The first two recommendations have serious limitations, as they would restrict scientists' sharing gender difference findings for several reasons, in a sense withholding data. The third recommendation would be acceptable if both similarities and differences were reported. Unquestionably, the fourth recommendation is praiseworthy. The fifth calls simply for researchers to do their job well by reporting samples fully, thus allowing generalization and replication. It is the sixth recommendation which provides the greatest potential, through the combination of two research modes.

In the L2 area, several of the recommendations are worth applying: reporting both similarities and differences by gender; reporting variances as well as means that show gender differences; describing samples in detail; and conducting multimethod research. If we follow these suggestions, it is possible that we will begin to untie the complex, interesting, and potentially important Gordian knot of gender differences in L2 learning styles and strategies. (see Note 2).

Notes

1. Rule-based practice might be considered analytic. However, in this study, females—who are often viewed as more global and less analytic than males—indulged in strategies related to rule-based practice significantly more often than did males. It could be that practical application of rules among many of these females was done in a global, patterned way rather than involving constant analysis of rules.

2. Selected ideas from parts of this article will appear in a volume entitled *Gender and English Language Teaching* (edited by Jane Sunderland for Prentice-Hall, 1993). Thanks to Lidia Woytak, Kenneth Kifer and two anonymous reviewers for their very helpful comments.

References

Associated Press. (1992, Aug. 1). Study shows brains differ in gay, heterosexual men: Anterior commissure area larger in homosexuals. *The Washington Post*, A2.

Baker, D. (1982). Differences in personality, attitude, and cognitive abilities found among biological, physical science, and nonscience majors. Paper presented at the annual meeting of the American Educational Research, New York.

Bardwick, J. (1971). *Psychology of women: A study of biocultural conflicts*. New York: Harper & Row.

Baumeister, R. F. (1988). Should we stop studying sex differences altogether? *American Psychologist, 43*, 1092-1095

Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162.

Bem, S. (1975). Sex role adaptability: One consequence of psychological androgyny. *Journal of Personality and Social Psychology, 31*, 634-643.

Birckbichler, D. W. (1984). The challenge of proficiency: Student characteristics. In G.A. Jarvis (Ed.), *The challenge for excellence in foreign language education* (pp. 47-78). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Block, J. H. (1973). Conception of sex roles: Some cross-cultural and longitudinal perspectives. *American Psychologist 28*, 512-526.

Boocock, S. S. (1990). *Sociology of education* (2nd ed.). Boston: Houghton Mifflin.

Brown, H. D. (1989). *A practical guide to language learning: A fifteen-week program of strategies for success*. NY: McGraw-Hill.

Brown, H. D. (1991). *Breaking the language barrier: Creating your own pathway to success*. Yarmouth, ME: Interculture Press.

Cahn, L. D. (1988). Sex and grade differences and learning rate in an intensive summer reading clinic. *Psychology in the Schools, 25(1)*, 84-91.

Caplan, P. J., MacPherson, G. M., & Tobin, P. (1985). Do sex-related differences in spatial abilities exist? A multilevel critique with new data. *American Psychologist, 40, 786-799*.

Chipman, S. (1988). Far too sexy a topic: Review of Hyde & Linn (Eds.), *The psychology of gender: Advance through meta-analysis*. *Educational Researcher, 17(3)*, 46-49.

Cohen, A. D. (1990). *Language learning: Insights for learners, teachers, and researchers*. New York: Newbury House/HarperCollins.

Cornett, C. E. (1983). *What you should know about teaching and learning styles*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Cox, D. & Waters, H. S. (1986). Sex differences in the use of organization strategies: A developmental analysis. *Journal of Experimental Child Psychology, 41(1),* 18-37.

Crawford, M. & Gentry, M. (Eds.). (1989). *Gender and thought: Psychological perspectives.* New York: Springer-Verlag.

Denmark, F., Russo, N. F., Frieze, I. H. & Sechzer, J. A. (1988). Guidelines for avoiding sexism in psychological research. *American Psychologist, 43, 582-585.*

Docking, R. A. & Thornton, J. A. (1979). Anxiety and the school experience. Research report, Australian Education Research and Development Committee, Canberra.

Dodds, D. (1991). Sex difference and the ACTFL oral proficiency interview. *Northeast Conference Newsletter* 34-36, 55-56.

Dunn, R. (1991). Do students from different cultures have different learning styles? *InterEd, 15,* 12-16.

Dyk, R. & Witkin, H. (1965). Family experiences related to the development of differentiation in children. *Child Development, 36,* 21-55

Eccles, J. S. (1989). Bringing young women to math and science. In M. Crawford & M. Gentry (Eds.), *Gender and thought: Psychological perspectives* (pp. 36-58). New York: Springer-Verlag.

Ehrman, M. E. & Oxford, R. L. (1988). Ants and grasshoppers, badgers and butterflies: Qualitative and quantitative exploration of adult language learning styles and strategies. Paper presented at the Symposium on Research Perspectives on Adult Language Learning and Acquisition, The Ohio State University, Columbus.

Ehrman, M. E. & Oxford, R. L. (1989). Effects of sex differences, career choice, and psychological type on adults' language learning strategies. *Modern Language Journal, 73,* 1-13.

Ehrman, M. E. & Oxford, R.L. (1990). Adult language learning styles and strategies in an intensive training setting. *Modern Language Journal, 74,* 311-327.

Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning, 32,* 367-391.

Ellis, G. & Sinclair, B. (1989). *Learning to learn English: A course in learn training.* Cambridge: Cambridge University Press.

Elias, M. (1992, Aug. 3). Difference seen in brains of gay men. *USA Today,* 8D.

Farhady, H. (1982). Measures of language proficiency from the learner's perspective. *TESOL Quarterly, 16,* 43-59.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Education Research, 62(1),* 61-84.

Fennema, E. & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and affective factors. *American Educational Research Journal, 14*, 51-57

Fishman, P. M. (1978). Interaction: The work women do. *Social Problems, 25, 397-406.*

Gage, N. L. & Berliner, D. C. (1975). *Educational psychology.* Chicago: Rand-McNally.

Galloway, V. & Labarca, A. (1991). From student to learner: Style, process, and strategy. In D.W. Birckbichler (Ed.), *New perspectives and new directions in foreign language education* (pp. 111-158). Lincolnwood, IL: National Textbook Co. and ACTFL.

Gannon, L., Luchetta, T., Rhodes, K., Pardie, L. & Segrist, D. (1992) Sex bias in psychological research: Progress or complacency? *American Psychologist, 47(3)*, 389-396.

Gass, S. & Varonis, E. (1986). Sex differences in NNS/NNS interaction. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 327-351). Rowley, MA: Newbury House.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge: Harvard University Press.

Good, T. L. & Brophy, J. E. (1986). *Educational psychology*, 3rd ed. New York: Longman.

Halpern, D. F. (1986). *Sex differences in cognitive abilities.* Hillsdale, NJ: Erlbaum.

Hansen, J. & Stanfield, C. W. (1981). The relationship of field dependent-independent cognitive styles to foreign language achievement. *Language Learning, 31*, 349-367.

Hansen, M. J. (1982). Spatial performance, activity preferences, and masculinity-femininity. Paper presented at the annual convention of the American Psychological Association, Washington, DC.

Hauck, L. S. (1985). Differences in information-mapping strategies in left- and right-brain learners. Paper presented at the annual meeting of the American Vocational Association, Atlanta, GA.

Hyde, J. (1981). How large are cognitive gender differences? *American Psychologist, 36*, 292-301.

Hyde, J. & Linn, M. C., (Eds.). (1986). *The psychology of gender: Advances through meta-analysis.* Baltimore: John Hopkins Press.

Jacklin, C. N. (1983). Boys and girls entering school. In M. Marland (Ed.), *Sex differentiation and schooling.* London: Heinemann.

Johnson, S. A. (1983)). Hemispheric specialization and learning style theory. Some considerations for the elementary teacher. Unpublished manuscript.

Keefe, J. W. (Ed.). (1979). *Student learning styles: Diagnosing and prescribing programs*. Reston, VA: National Association Of Secondary School Principals.

Kramarae, C. (1981). *Women and men speaking*. Rowley, MA: Newbury House.

Lakoff, R. (1975). *Language and women's place*. New York: Harper & Row.

Larsen-Freeman, D. & Long, M. (1991). *An introduction to second language acquisition research*. London: Longman.

Lawrence, G. (1984). A synthesis of learning style research involving the MBTI. *Journal of Psychological Type, 8*, 2-15.

Leaver, B. L. (1991, Nov. 24). Personal communication.

Leaver, B. L. (1986). Hemisphericity of the brain and foreign language teaching. *Folia Slavica, 8*, 76-90.

Linn, M. C. & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479-1798.

Lusk, E. J. & Wright, H. (1981). Difference in sex and curricula on learning the Group Embedded Figures Test. *Perceptual and Motor Skills, 53(1)*, 8-10.

Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford: Stanford University Press.

Mansnerus, L. (1989, Aug. 6). SAT separates girls from boys. *New York Times*, 4A, 27-28.

Mendelsohn, D. (1984). There *are* strategies for listening. *TEAL Occasional Papers, 8*, 63-76.

Mills, C. (1980). Sex-role-related personality correlates of intellectual abilities in adolescents. *Roeper Review, 2(3)*, 29-31.

Moirs, A. & Jessel, D. (1991). *Brain sex: The real difference between men and women*. New York: Stuart/Carol.

Myers, I. B. & McCaulley, M. H. (1990). *Manual: A guide to the development and use of the Myers-Briggs Type Indictor*. Palo Alto: Consulting Psychologists Press.

Naiman, N., Frohlich, M. & Todesco, A. (1975). The good second language learner. *TESOL Talk, 6(1)*, 58-75.

Nyikos, M. (1987). *The effect of color and imagery as mnemonic strategies on learning and retention of lexical items in German*. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *Modern Language Journal, 74(3)*, 273-287.

O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press

Orlorfsky, J. L. & Stake, J. E. (1981). Psychological masculinity and femininity: Relationship to striving and self-concept in the achievement and interpersonal domains. *Psychology of Women Quarterly, 6(2)*, 218-233.

Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implication for strategy training. *System, 17*, 235-247.

Oxford, R. L. (1990a). Language learning strategies and beyond: A look at strategies in the context of styles. In S. S. Magnan (Ed.), *Shifting the instructional focus to the learner* (pp. 35-55). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Oxford, R. L. (1990b). *Language learning strategies: What every teacher should know*. New York: Newbury House / HarperCollins.

Oxford, R. L. (1990c). Styles, strategies, and aptitude: Connections for language learning. In T. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 67-125). Englewood Cliffs, NJ: Prentice-Hall.

Oxford, R. L. & Crookall, D. (1989). Language learning strategies: Methods, findings, and instructional implications. *Modern Language Journal, 73*, 404-419.

Oxford, R. L. & Ehrman, M. E. (1988). Psychological type and adult language learning strategies: A pilot study. *Journal of Psychological Type, 16*, 22-32.

Oxford, R. L. & Lavin, R. Z. (1991). Teacher-student "style wars" in the language classroom: Research insight and suggestion. *Bulletin of the Association of Department of Foreign Languages, 23(2)*, 38-45.

Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal, 73*, 219-300.

Oxford, R. L., Nyikos, M. & Ehrman, M. (1988). *Vive la difference?* Reflections on sex difference in use of language learning strategies. *Foreign Language Annuals, 21(4)*, 321-329.

Oxford, R. L., Park-Oh, Y., Ito, S. & Sumrall, M. (in press). Learning Japanese by satellite: What influences student achievement? *System, 21(1)*.

Parry, T. S. (1984). *The relationship of selected dimensions of learner cognitive style, aptitude, and general intelligence factors to selected foreign language proficiency tasks of second-year students of Spanish at the secondary level*. Unpublished doctoral dissertation, Ohio State University, Columbus.

Politzer, R. L. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievement. *Studies in Second Language Acquisition, 6*, 54-68.

Riger, S. (1992). Epistemological debates, feminist voices: Science, social values, and the study of women. *American Psychologist, 47(6)*, 730-740.

Rigney, J. W. (1978). Learning strategies: A theoretical perspective. In H. F. O'Neil (Ed.), *Learning strategies* (pp. 165-285). New York: Academic Press.

Roberts, T. (1979). Reflection-impulsivity and reading ability in seven-year-old children. *British Journal of Educational Psychology, 49(3)*, 311-315.

Rothblum, E. E. (1988). More on reporting sex differences. *American Psychologist, 43*, 1095.

Rubin J. (1975). What the "good language learner" can teach us. *TESOL Quarterly, 9(1)*, 41-51.

Rubin, J. Z., Provenzane, F. J. & Luria, Z. (1974). The eye of the beholder: Parents' views on sex of newborns. *American Journal of Orthopsychiatry, 44*, 512-519.

Schmeck, R. R. (1983). Learning styles of college students. In R. F. Dillion & R. R. Schmeck (Eds.), *Individual differences in cognition*, Vol. 1 (pp. 233-279). New York: Academic Press.

Serbin, L. A. & O'Leary, K. D. (1975). How nursery schools teach girls to shut up. *Psychology Today*, 57-58, 102-103.

Sherman, J. A. (1978). *Sex-related cognitive differences: An essay on theory and evidence*. Springfield, IL: Thomas.

Shipman, S. & Shipman, V. C. (1985). Cognitive styles: Some conceptual, methodological, and applied issues. In E. Gordon (Ed.), *Review of Research in Education*, Vol. 12 (pp. 229-291). Washington, DC: American Education Research Association.

Sigel, I. E., Jarman, P. D. & Hanesian, H. (1967). Styles of categorization and their intellectual and personality correlates in young children. *Human Development, 10*, 1-17.

Slavin, R.E. (1988). *Educational psychology: Theory into practice* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Smith, C. & Lloyd, B. (1978). Maternal behavior and perceived sex of infant: Revisited. *Child Development, 49*, 1263-1265.

Springer, S. & Deutsch, G. (1989). *Left brain, right brain*. New York: Freeman.

Sunderland, J. (in press). *Gender and English language teaching*. Englewood Cliffs, NJ: Prentice-Hall.

Tannen, D. (1986). *That's not what I meant!* New York: Morrow.

Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York: Ballentine.

Thorne, B., Kramarae, C. & Henley, N. (1983). *Language, gender and society*. Rowley, MA: Newbury House.

Torrance, E. P. (1962). Developing creative thinking through school experience. In S. J. Pranes & H. P. Harding (Eds.), *A source book for creative thinking*. New York: Scribner.

Tran, T. V. (1988). Sex differences in English language acculturation and learning strategies among Vietnamese adults age 40 and over in the United States. *Sex Roles, 19(11-12)*, 747-758.

Tyler, L. (1965). *The psychology of human difference*. New York: Appleton-Century-Crofts.

Vann, R. & Abraham, R. (1989). Strategies of unsuccessful language learners. Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, San Francisco.

Wandersee, J. H. (1988). Ways students read texts. *Journal of Research Science Teaching, 25(1)*, 69-84.

Weiner, N. C. & Robinson, S. E. (1986). Cognitive abilities, personality, and gender differences in math achievement of gifted adolescents. *Gifted Child Quarterly, 30(2)*, 83-87.

Weinstein, C., Goetz, E. T. & Alexander, P. A. (Eds.). (1988). *Learning and study strategies: Issues in assessment, instruction, and evaluation*. New York: Academic Press.

Wenden, A. (1991). *Learner strategies for learner autonomy: Planning and implementing learner training for language learners*. Englewood Cliffs, NJ: Prentice-Hall.

Wenden, A. & Rubin, J. (Eds.). (1978). *Learner strategies for language learning*. Englewood Cliffs, NJ: Prentice-Hall.

Willing, K. (1988). *Learning styles in adult migrant education*. Adelaide, S. Australia: National Curriculum Research Council.

Witkin, H. A. & Berry, J. W. (1975). Psychological differentiation in cross-cultural perspective. *Journal of Cross-Cultural Psychology, 6*, 4-87.

Author

REBECCA L. OXFORD, Professor, Department of English and Area of Curriculum & Instruction, The University of Alabama, 204 Graves Hall, Tuscaloosa, AL 35487. Specializations: language learning styles and strategies, language testing, language teaching methodology, psychological aspects of development.

Applied Language Learning
1993, Vol. 4, Nos. 1 & 2, pp. 95-130

Sheltered Initiation Language Learning *

Zev bar-Lev
San Diego State University

This paper reports on a multi-language project for development of foreign-language curriculum, and the teaching method which has developed from it, which aims at giving confident, continuous, and creative speaking abilities. The method is represented primarily in a set of "mini-courses," each being a short introduction to a given language.

The main features of the method are: (1) Students first learn to concatenate words as they learn them, thus creating their own discourses, rather than first learning pre-fabricated whole conversations. (2) The distraction of grammatical inflection is deferred until after the initial "pre-language" stage, which has the power of a pidgin or telegraphic stage of L1 acquisition, but none of the deviance of these simplified forms of language. (3) Grammatical inflection is subsequently introduced "asymmetrically," one category at a time, for maximal absorption into the developing fluency. (4) Comprehension skills of skim-reading and "skim-listening" are taught with special exercises.

Scope of the Paper

This 7-year multi-language project for creating an innovative foreign-language curriculum and the teaching method which has evolved as a result aim primarily at developing speaking abilities, focusing on initial instruction. The method has been described with respect to Hebrew (bar-Lev, 1991a-b), Arabic (bar-Lev, 1991c), and Chinese (bar-Lev, 1991d), and has been cited for its rapid results (Groesbeck, 1988; Celce-Murcia, 1991).

* This article and the development behind it were funded in part by the National Language Resource Center of San Diego State University under a grant from the US Department of Education, Center for International Education.

© 1993, Zev bar-Lev

General Background

The sheltered initiation language learning method (SILL) is represented primarily in a set of mini-courses, each a short self-contained introduction to a language, representing twenty college-level class hours or fewer. Fifteen languages have been taught in mini-courses, most of them to seven or eight generations of students.

The method has been used for up to two years of regular curriculum, as the basis for therapies for students who wanted to acquire speaking abilities in languages that they had studied for several years but could not speak, and as models for teaching speaking more effectively at all levels.

Following Robert W. Blair's suggestion in his review of the method (in Celce-Murcia, 1991) I will call it *Sheltered Initiation Language Learning (SILL)*. There is only a partial similarity to sheltered English, but the general concept of "sheltering," that is, protecting the student from the mass of linguistic detail in the language as a whole, is appropriate, since the method aims to achieve its goals within a deliberately limited (although generic) vocabulary and grammar.

The method asks how much vocabulary and grammar various learners (strong versus weak; future professors of the language versus students merely completing a language requirement; and tourists with limited time for study) can acquire for productive, spontaneous use. The project has been guided throughout its seven years by a no-holds-barred approach to the goal of making study of a foreign language as effective as possible vis-à-vis ultimate goals, in particular as regards speaking, and beginning in the initial weeks of study. Details that are not relevant to immediate development of speaking skills have been deferred or ignored.

Although the first weeks and months of curriculum are made easy, language *skills* are not deferred in the least; on the contrary, the project has developed a direct link to several realistic language skills. This has led to the concept of "language in real time," allowing considerably more realism than is typical in the foreign language classroom; thus the study instructions for mini-course students might appropriately be addressed to advanced learners, when they are about to confront native speakers in real situations.

The final form of the curriculum, after 7 years of development with several hundred students in a variety of languages, is radical as compared with current teaching theory and especially practice. But there is reason to suspect a need for radical solutions (see Simon, 1988); at least there is no reason to forego the search for them, nor can one demand that any curricular innovation be justified in advance by experimental data, since a controlled experiment can be set up only after the innovation has been formulated. There is every reason to expect innovations to take considerable time to define, try out, and refine,

whether they are intended for controlled experiments or for the classroom alone.

Methodology of this Report

The exposition of SILL will follow a topical order, explaining its three main characteristics: *pre-language*; *asymmetric sequencing*; and *realism of skills*. These features of the method will be described as they arose in response to specific goals and problems. The reason for this ordering is that the methodology of this report is more qualitative than quantitative.

Qualitative studies have become increasingly popular in methodological research. Perusal of *Foreign Language Annals*, for example, reveals a number of reports on new methods and programs without quantitative validation (García, 1989; Oxford et al, 1990; Poritsky, 1990). One can even find subjective evaluations within these reports. This may include a random epithet, as well as surveys in which large numbers of subjective reactions may be summarized. In García (1989, p. 366) for example, we find the assertion—given without statistical or bibliographical support—that "improved proficiency will occur as students become aware of the ways L1 speakers relate to each other in authentic situations" (an example of the broad methodological assumptions that teachers make, ideally but not invariably with relevance to broader psychological literature). In Oxford et al (1990, p. 202), we find unquantified and subjective statements like: "Cohen receives many positive and often ebullient comments from students. Many learners report . . . Learners say . . ." Just as the clinical psychologist would be unable to practice simply on the basis of concepts proved by controlled experiments, so too the language teacher would be unable to teach simply on the basis of controlled experiments in language teaching. Foreign language teaching itself is, after all, a "clinical therapy," which cannot always wait for controlled proofs, but must often proceed on its own. Individual successes, although capable of misinterpretation, are what usually drive the teacher's day-to-day methodological decisions—apart from basic methodological assumptions. Much innovation in foreign language teaching requires shifts in broad goals, and the present innovations are no exception. That a given exercise works is often logically subordinate to what it accomplishes, a consideration which teachers trained under older goals may not appreciate.

In 7 years of the project, I have only a few times had groups large enough for controlled experiments, and these I would prefer to call pilot experiments because they raised many methodological questions. Much

of my effort has necessarily been spent formulating innovations, and even where they were immediately successful, refining them with successive classes. Offsetting this inability to do many controlled experiments, I have had the advantage of teaching mostly classes of what might be called "limited effort" students: The classes have always been optional—in the extreme case, working adults meeting once a week. Even in a large Hebrew program, students had the option of completing the language requirement with easier languages, and many of them did so while studying Hebrew for one or two semesters on their own.

The advantage of these limited-effort frameworks was that they created pressure for successful curriculum development: Any little curricular mistake, such as putting two similar words in the same lesson, would have had clear and unwelcome consequences, causing the student to make immediate, repeated, often irremediable errors—or even to drop out of the course. Whether or not the present form of the curriculum is optimal, it has passed the gauntlet of exceptionally demanding students.

In any case, the present survey of the project is more qualitative than quantitative: While a few pilot experiments will be referred to for results that have been investigated quantitatively, the primary motivation offered for many of the methodological conclusions outlined will be the etiology of their development, and the response that they seemed (albeit without controls) to elicit among actual students. We thus survey the main features of the project as they emerged and developed.

Pre-Language

vis-à-vis Natural Approach

The method of SILL parallels the natural approach (Krashen & Terrell, 1983) in proclaiming the goal of acquisition as primary, and insisting that acquisition be attestable in spontaneous speaking abilities. Further, the natural approach insists on naturalness of acquisition to the extent of precluding any sort of interference. But the classroom is an inherent invasion of the natural process, and bringing "naturalness" into the classroom must involve adjustments, if not equivocations. Many readers of the literature of the natural approach are surprised to hear that two out of five class hours are devoted to grammar study, in actual implementations of the method, but Tracy Terrell candidly admitted that publishers determine grammar sequencing for his textbooks, as for others (see also Anonymous, 1990).

VanPatten (1992) has argued the need, in communicative language teaching, for *structured input* and *structured output*—a grammatical orientation to specific activities, including activities in which production of output on the part of the student is the immediate goal, rather than the long-term, natural result of acquisition via input alone. VanPatten has also argued for the importance of beginning with single sentences rather than entire discourses, and of teaching one grammar point at a time. All of these ideas have been utilized in my project for several years (although without VanPatten's extensive research base). While some other major curricular concepts in this project may for the present remain unorthodox, in a fast-moving field such as language pedagogy, ideas may be of interest even if they have emerged from non-traditional directions and through trial-and-error curriculum development.

Inflection and Telegraphic Pidgin

The SILL program was inspired by a challenge from a student of psycholinguistics. Discussing the telegraphic stage (in which the child first strings words together into full sentences, although without grammatical inflections), I would ask whether any foreign-language teacher ever allowed students to communicate this way. This stage, after all, offers the possibility of complex communication, without the complexities of inflectional morphology which many students find difficult, so it is a good example of the genius of natural acquisition. A student challenged me to offer foreign language studies along these lines, and a handful of students embarked the next semester on individual studies, using vocabularies that I had set up over the winter break, beginning to study what was in effect a telegraphic pidgin of their languages of choice.

In surveying students I have found they are often overwhelmed by grammatical inflection. Many remember conjugation more than any other experience in Spanish and French, and similarly declension in German and Russian. They usually report having studied more tenses than they are able to use in sentences in sequence spontaneously. For example, one relatively sophisticated class (linguistics students) yielded the following profile: The average student in this class had studied Spanish for 3 high school years, earning a grade of B-plus, with teachers and books both described as "good." This average student could read authentic materials with a dictionary, and recalled studying an average of 8 tenses and 27 irregular verbs, but would expect to be placed in the second college semester—and would be able, by self-report, to speak no more than three sentences spontaneously in sequence.

In less commonly taught (LCT) languages the majority of students never achieve spontaneous speaking abilities; I have known graduate students and even teaching assistants in Russian and Hebrew who could not carry on a simple conversation—although all of them could conjugate magnificently on paper and in oral exercises. I have found through surveys that students overwhelmingly attribute their lack of fluency to grammatical inflection: Gender and conjugation force a mid-sentence calculation that brings their attempts at speaking to a halt.

Students in my surveys overwhelmingly considered speaking to be their primary skill goal, even in Hebrew, which they often take primarily for religious reasons. The working assumption of the project is that initial confidence is crucial in this most important aspect of language. Notwithstanding traditionalists' fears and apart from its inherent benefits (especially for weaker students), the method prepares students to deal with the difficulties of language at later stages. The efficiency of the resulting curriculum may even allow later coverage of the same points of grammar.

Lexical Initiation

Because of the focus on lexicon and the telegraphic stage, the project initially resembled the lexical learning project set up by John Schumann (Cushing et al, undated). While such a starting point may not be appealing from the perspective of natural acquisition theory, it does make sense in the perspective of pidginization theory (Schumann, 1986), and in terms of the problem of grammatical inflection. For a while I focused on pidginization in acquisition, anticipating that a lexical base could serve as the starting point for natural acquisition. Like Schumann, I was impressed with the power of the lexical initiation: My students, like Schumanns', were all able to speak with notable ease and "fluency." I assumed that, although telegraphic speech may not be the ultimate goal, it gives such remarkable initial confidence and fluency that it must somehow be incorporated into the ideal teaching method.

The idea of starting with lexical items rather than full sentences, or even full conversations, is unusual if not heretical. It is not that context is undervalued, but rather that, with a minimum of examples, students are expected to create their own sentences from the beginning. I suggest that such pre-fabs as sample conversations are counter-productive. Students must ultimately learn to build up their own discourses brick-by-brick; pre-fabs will not help them.

In a pilot experiment, I compared lexical and sentential initiation in Arabic. Students were taught a mini-lesson which presented words with a sample sentence for the experimental group, while students in

the control group were given full sentences to memorize. The test consisted of an additional lesson which taught two new words and a sentence using each of them, followed by a time-limited English-to-Arabic translation exercise. The total correct number of words for each student was counted. The control group averaged 41 percent, the experimental group 68 percent. Many more answers in the control group were incomplete, showing the lesser fluency reflected in these averages; in other words, if complete sentences had been required for correctness, rather than just individual words, the results would have been more disparate.

It was also interesting to observe attitudinal differences between the groups: The control (sentential) group began stronger, since they were able to speak in sentences from the beginning. The experimental group started less confidently, since they had to create each sentence on their own. The tide turned after a few minutes, as I would suggest it must in any similar comparison (although not necessarily so soon).

Students of the mini-courses must start by struggling to create their own sentences, achieving fluency only gradually. Some students may take several class hours before they can say sentences of more than two words with great confidence and spontaneity. But in a sentential initiation, the learners become discouraged later, as the weight of memorizing long sentences becomes burdensome and they begin to feel bored and incompetent. Learning conversations must inevitably involve rote memorization of ever increasing numbers of long sentences.

More specific discussion of fluency will be left for later, after we have dealt with the important problem of grammatical inflection and accuracy. In the following section we will prepare the way for the discussion of accuracy by noting how the telegraphic pidgin of the pre-language stage diverges from the structure of a pidgin (and telegraphic stage).

One Form per Word

The initial stage of the mini-course is a telegraphic pidgin as described above; with the one major modification that since the goal is to ultimately achieve as much accuracy as possible, this initial "pre-language" stage is restricted so as to fall within the limits of grammatical acceptability. The pre-language stage is limited to one form per word (e.g., the masculine of adjectives, the 3rd-person singular present of verbs). All words that would grammatically conflict with these forms (feminine nouns, personal pronouns) are deferred. It is thus a purposeful, even contrived, selection of the target language.

The goal is to provide the learner with the psychological ease and communicative power of the telegraphic pidgin, but without any of the deviance that normally accompanies this stage of first and second

language acquisition because of the perceived dangers of fossilization, as well as to allow even traditional teachers to use it. While some communicative range is obviously lost with this choice, the possibility of integration into an ordinary language classroom is a good tradeoff, apart from the still uncertain dangers of fossilization. Practice showed that while communicative range was temporarily restricted by this move, learner fluency as such could be developed equally well.

Children and other natural acquirers take the approach of making errors while acquiring grammatical details gradually—errors that are not likely to be tolerated in ordinary classrooms. SILL took a different direction as it developed, refining the initial stage so that indeed it would contain no inflectional errors, simply by deferring, for Spanish, feminine nouns and personal pronouns like *yo*, thus providing the initial confidence of a pidgin without its inaccuracy. Thus, not only were multiple forms of single words deferred, but also any word that might clash with the chosen form: where feminine adjectives were excluded from the pre-language stage, so too were feminine nouns; where the 3rd-person was the first verb form taught, other personal pronouns were deferred. The form chosen as initial depended on the language. The masculine is presented first in the Romance languages and Russian, the feminine in German and Bulgarian; the 3rd-person singular is chosen in Romance, but the 1st-person in German, and the imperative in Arabic. (The reasons for these various choices are too complex for the present discussion.)

For many languages, this choice meant an exclusive focus on narrative and descriptive, as opposed to dialogic ("I-you"), speech. But a narrative-descriptive focus was in any case increasingly being seen as the best arena for development of discourse abilities: Dialogue was seen to be not only grammatically complicated by the "I-you" switch (apart from grammatical inflection), but also more limited in the kinds of memory demands it makes, and therefore not as useful for developing speaking goals. Thus, when SILL students learn to tell time, they do so with narrative: "It's three o'clock. Mother is sleeping . . ."

As the program has developed, I have tried shortening the pre-language stage gradually, to enhance the quality of grammatical acquisition. The effect is two-fold: First, allowing more time for each grammatical category enhances its acquisition; second, giving learners slightly less confidence with the pre-language seems to motivate them for grammar as such, not at all a negative outcome. The pre-language stage is now a mere two hours long, and more emphasis is put on the appropriate presentation of grammar, with no apparent sacrifice of fluency.

Asymmetric Sequencing

Grammar

Even in pre-language, the students' grammatical accuracy was more impressive than might be the case in a true pidgin (although not because of any effort on their part): By teaching a single form for each word, and more specifically choosing for most languages the 3rd-person singular of verbs rather than the infinitive, I was able to ensure that a large proportion of their inflectional endings (30 to 50 percent, depending on the topic and language) were correct in context. This was accomplished without effort or even an awareness on their part that they were using endings, but simply because I had selected the most common single word form. (With the infinitive as the first verb form taught, essentially zero percent of sentences can be correct.) Thus, for example, in Spanish, a student could theoretically say a sentence like *El fin del siglo estimula el desarrollo del pueblo* without knowing anything about gender or conjugation. (This sentence happens to be beyond the vocabulary taught in the Spanish pre-language.)

Of course, one cannot continue for long with this partial approach. Sooner or later, one must confront the problems of gender and conjugation. While coming from a naturalistic philosophy of language teaching, I am not optimistic that errors will correct themselves in natural environments; while some may work themselves out, or simply not recur, others may remain with a student simply because he is proud of his early "creations" in the language. Natural acquisition is no guarantee of accuracy: Natural acquirers of Hebrew often make horrendous mistakes with the special construction for *have*. As a teacher of LCT languages, I cannot believe that natural acquisition will ultimately create a satisfactory version of the language.

To the extent that we aim for accuracy in speaking, I believe that we must do so via curricular sequencing (as well as exercises), rather than relying on natural acquisition. Recognizing the impossibility of total accuracy may help us set our sights realistically, but I suggest that we will be better off investing our efforts in developing better presentations and exercises, rather than hoping for some unlikely renaissance of grammatical awareness among our students.

It cannot be too strongly emphasized that memorization of whole phrases and sentences, although used in the audio-lingual method, does not seem like a good general strategy for achieving accuracy. For example in Spanish students who learn *me llamo* (I call myself) as a fixed phrase often turn it into *me llamo es* ("I-call-myself is") as soon as they learn *es* (is). This familiar fact of acquisition must be added to the general reasons for a lexical initiation. Since memorization fails to teach

either sentence patterns or idiomatic constructions, it must be regarded as a failure except for fixed expressions such as *Hello*. For this reason, no doubt, many teachers and textbooks no longer obligate their students to memorize dialogues; but the dialogues lose their power as a source of fluency unless memorized.

In the SILL project I have been able to incorporate grammatical accuracy gradually, ultimately with no diminution of fluency. Any slowing of fluency, even for one or two students, was a sign of a bug in the curriculum, which I attempted to solve by redesigning the curriculum for subsequent groups. This sort of gauntlet is not a controlled experiment, but it does parallel the way most teachers teach and develop curriculum.

Defining the inflectional goals for a given length of study was a crucial issue. Early exposure to multiple tenses, for example, was found to have a harmful effect on most students: They would imitate the past tense or even confuse it with the present, as well as forming it incorrectly. While a fuller exposure to a target language environment might have influenced some of this, there seemed to be no possibility of solving the problem within the limits of the classroom. While in the natural approach one might assume that these inaccuracies would work themselves out gradually, I had already seen mistakes persevere over intolerably long periods in natural acquisition of Hebrew in Israel. Certainly no amount of input in the classroom could guarantee a better outcome.

Over the years, partly with the help of expanding the number of languages taught and the comparison that such a multi-language effort allows, sequencing has been found to be a powerful tool for successful acquisition. By trial and error with successive groups, I was forced to develop a sequencing principle, *asymmetric sequencing*, which turned out to be effective for inflectional morphology, for phonetics and orthography, and even for vocabulary.

Asymmetry emerges naturally from the one-form-per-word structure of the pre-language: Since only masculines were presented first, feminines would have to be presented as a second category. This differs from the traditional presentation, in which the problem of gender confronts the student from the beginning. Where Tracy Terrell's *Dos Mundos* (1990) begins with *you* (familiar) and *I* forms at the same time, the Spanish mini-course begins with the *he/she* form, and later introduces the *I* form.

In communicative language teaching, sequencing is supposed to occur naturally, as determined automatically by the authentic materials chosen (Rutherford, 1987). But if VanPatten's (1992) suggestion is followed, the issue of sequencing could resume its traditional prominence. At present, grammar sequence is determined rigorously by publishers (Anonymous, 1990): Even teachers who sequence naturally

are not necessarily unaware of the grammatical and lexical dimensions of the resulting sequencing, and in classroom practice a commitment to the natural approach may often degenerate into mere contextualization of grammatical forms taught according to schedule.

Asymmetric presentation is also more consistent with the notion of "markedness" in language: The Spanish 1st-person is more "marked" (less neutral) and therefore should be learned as special. While other languages do not have freely occurring stems as English does, they do have more neutral and less neutral forms; for example, Spanish *viene* is more neutral than *vengo*. (The maximally neutral verb form, the infinitive, is not chosen because it is less frequently grammatical in discourse: *yo viene* may not sound much better than *yo venir*, but it was felt to be closer to the correct form and therefore a better starting point.)

On a practical level, the overuse of the unmarked form, although no less an error than overuse of the marked form, is less confusing to the native listener: *Yo viene* is not to be encouraged, but it is less confusing than *Él vengo*, and if errors are inevitable, surely a curriculum will be more highly valued if, *aliis paribus*, it allows more optimal errors than non-optimal ones. Teachers do not like to face the question of optimal errors, but this is like jumping for the moon: Even when considering accuracy, is it not relevant to ask what a B or C student can use at the end of a given course, since B's and C's together constitute the majority of marks?

Similarly in conjugation, for example in the Romance languages, where the ordinary textbook introduces six forms at once (even where some of them are marginal or identical), it was natural to follow up the 3rd-person singular present in the mini-course with the main differing form (*I* for Spanish and Italian, *you* for French), as well as to ignore the familiar *you* form until after the mini-course. This is similar to VanPatten's cited proposal to teach one thing at a time.

Use of the polite rather than the familiar may seem strange among students in a classroom. (Role-play must be used to put students into the appropriate relationships, but it is otherwise necessary anyway.) Again the notion of optimal error is appropriate. I would rather have my students overuse the polite, rather than the familiar, since it may be amusing but not rude or dangerous to do so.

Some students catch on immediately to a new grammatical category when it is presented asymmetrically. Interestingly enough, even stem changes have not seemed to present a problem for learners. A given learner might not immediately succeed with *viene* > *vengo*, for example, but no great frustration was engendered either way in the mini-courses. On the other hand, presenting the *-oy* ending for *I* in Spanish along with the *-o* ending brought some students to a dead stop in acquiring morphology. This could be a function of the specific groups that were studying at the time that this issue came up (although they

were more sophisticated grammatically), but there is no urgency about testing this with controlled experiments: The important thing is that the principle of asymmetry allows inclusion of stem changes, as I subsequently found with Italian *va > vado* and Russian *prinosit > prinoshu*, partly because the regularized forms *prinosu* or *prinosyu* approximate the correct forms. However, two or three different types at once can be confusing.

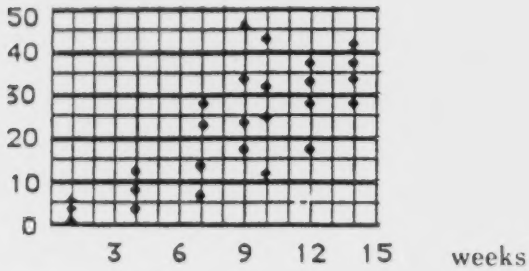
Fluency

While it would be nice to provide statistics in the form of ACTFL proficiency levels for broad numbers of students, this is impossible as well as inappropriate, in view of the fact that SILL more or less ignores trivial topics and vocabulary which are a focus of the proficiency levels as currently defined. In terms of proficiency levels, SILL students routinely achieve some aspects of the intermediate level and beyond in the mini-course although in doing so they bypass some aspects of the novice level. Thus, they cannot list days of the week or months of the year or catalogue types of clothing, but they readily speak in discourses several sentences long, and can narrate events in the past (though using the present tense to do so, as is allowed in many languages). They are weak (as well as not very interested) in discussion of what clothing they are wearing, and may not be able to describe their room, but they are able to state and explain preferences, for example for a big school or a small school, and can describe career goals (although not for a wide variety of careers) and discuss other matters of preference and opinion.

Although proficiency is not usually defined in a quantitative way, the spontaneous productions of various classes of SILL students have been observed quantitatively, in terms of sentence length and discourse length. *Figure 1* is a survey of the words-per-discourse of four students of Japanese over a 15-week mini-course; these data are typical for other classes and languages as well.

Figure 1

words per discourse



Of the hundreds who had studied in mini-courses (excluding one-hour-per-week adults), only a few special cases have emerged without general speaking confidence. For example, of 60 students of 1st-semester German, only one was unable to speak into the tape recorder for the mini-course final, although he had been speaking spontaneously in front of the class. All of this is not a bad achievement for these mini-courses of 20 class hours and less.

Students have from the beginning of the multi-language project routinely achieved fluency to about 30 words per minute by the end of the mini-course. Although this result is perhaps no more than a quarter of native speech rates, it is achieved in a continuous manner. Slower students (including some but not all of the learning disabled) may achieve rates of only 15 wpm, with noticeably more limited vocabulary, but still exhibiting continuity (without hesitating, hemming and hawing, self-correction, or English insertions). Others develop rates in the range of 50 wpm. Students are not limited to fixed dialogues, since none are given, or even to sentences, since only one sample per pattern is given.

In the first semester in which I taught a full Hebrew course, the final exam produced spoken answers averaging 38 words per discourse, with several answers above 80 words. (Because of Hebrew syntax, there are fewer words per sentence than in Japanese, for example, "the book is in the house" is a two-word sentence in Hebrew.)

Long-term retention is high even without study : A year after a mini-course, randomly selected students have been able to respond readily to the challenge to speak, achieving an average rate of 25 to 30 wpm, with the ability to say several sentences of their own creation in sequence, although some of the vocabulary had been forgotten after this long a lapse. Those 20 to 30 students who were followed in subsequent coursework reported retaining benefits of the confident start, at least on an emotional level, for a year or more, although in some instances these benefits eventually wore off under the pressure of ordinary coursework.

Students of Hebrew, more than any other language, have used the mini-courses to initiate fluency in a language that they had already studied for reading and grammar, so I have been able to observe that students succeed far better when they begin with the mini-course. In Hebrew, newcomers regularly produce 200 to 400 percent as much discourse as "recycled" students, in spontaneous speaking and in rapid writing and ordinary writing; in other languages similar differences have been noted. While some of the recycling students bring greater grammatical knowledge, which sometimes translates into greater accuracy, most often it does not; the main difference between newcomers and old-timers is often that the latter are more self-conscious. Even an old-timer with good grammatical command must revert to inaccuracy while moving to new levels of communicative range.

Accuracy has been followed as well, for some classes. 75 percent of a Russian class (all the A and B students) were able to speak spontaneously with the three personal endings of the mini-course. They reported an instant of hesitation while forming the new endings, during the first day that they encountered each one, but I could not hear any difference in pace, much less an actual hesitation. They occasionally produced a regularized incorrect form (e.g. *prinosyu* for *prinoshu*) over a few days. The infinitive was more difficult: Even in writing, any use of it, not to mention command of its complex forms, was sporadic.

In natural approach classrooms, students exhibit habits of silence, reliance on one-word sentences, habits of hesitation and self-correction, and habits of ending all sentences, for instance, with *¿como se dice?* All these tactics are allowed in the natural approach, but not in SILL. Grammatical errors such as *yo viene* which occurred in early versions of SILL have become less common with recent versions. Worse, mistakes occur in natural approach which are typical of students of grammar-based methods (*Usted voy* is generated in the natural approach from the prominence of the first person in input, rather than from its place at the head of a paradigm). In accordance with the idea of sheltering (although this term traditionally refers to input rather than output), I feel that restricting students' grammatical ambitions allows a higher degree of accuracy within the deliberately limited grammatical range. It has remained a rule throughout 7 years of intensive work on details of grammar (as well as phonetics) that no student should be banished explicitly or implicitly from study of a language because of an inability to acquire a particular grammatical detail, even an important one. Teachers who believe that students learn only what they are forced to learn may be relieved to hear that this is not the case.

A Pilot Experiment

A pilot experiment compared SILL with an ordinary communicatively-oriented class (using Moeller & Liedloff, *Deutsch heute*, 1984). The main dimensions for comparison were fluency and quality, as measured indirectly in a timed test, in which students responded in rapid writing to communicative questions, such as "Who are you?" (I am grateful to Mary Wauchope of San Diego State University for her participation in this experiment.)

Fluency was measured by the total number of German words in the answers, and quality was measured by holistic grading, with both of the classroom teachers marking all papers. Identifying marks were removed and papers were evaluated on the scale which I use in SILL (see *Figure 2*). The results were totaled for each student and averaged for each teacher and class, to produce the qualitative assessments in *Figure 3*. Each teacher preferred her own students by a similar margin (control

1.3, experimental 1.1). This may relate to the variation in the questions which has been noted, as well as to biases, since a teacher can often identify his own students' work without identifying marks. I suggest that no significant variation can be seen here.

Figure 2

4: very interesting responses
3: fully adequate responses
2: minimally adequate or almost adequate responses
1: inadequate responses
0: no responses

Figure 3

	TEACHER	
	experimental	control
CLASS:		
experimental:	10.25	10.9
control:	9.62	12.3

No attempt was made to assess the grammatical quality of the rapid writing responses separately. It was felt, on communicative grounds, that grammatical quality should not be measured without a coordinated consideration of communicative daring, and the latter seemed hard to measure at all. Much faulty grammar was noted throughout responses in both groups. Some impressions on individual answers might be noted, at least as a way to show the complexity of the whole question of communicative quality. The control group's most interesting answers to "*Who are you?*" were along the lines of *Ich bin ernst und konservativ. Ich bin progressiv.* These appear interesting, except that they are, word for word, among the literal answers provided by the textbook, and which the students had specifically practiced.

By contrast, the most interesting answers to this question in the experimental group were, *Ich bin eine Studentin und Mutter* or, *Ich bin ein Student, ich lerne deutsch.* These are objectively more prosaic answers, as well as grammatically imperfect.

These specific answers are not provided in the mini-course; no such question is asked. As spontaneous answers to unexpected questions, I suggest that they are more novel psychologically than the control group's answers. Unfortunately, it seems impossible to quantify this important question in a suitably information-theoretic way We cannot easily answer the question of how much originality is demonstrated with reference to a given curriculum. This issue is

relevant not only to assessment, but also to student enjoyment and accomplishment. If independence is the goal, it is obviously fostered less by a curriculum that provides whole answers, no matter how interesting they are.

The design of the test merits further discussion. The original questions proposed for the test are shown in *Figure 4*. When I had taught the German mini-course to two classes of similar size, I used fifteen such questions as the oral final after three weeks of instruction, and found that all students were able to respond, comprehensibly and individually, in one to four sentences. The control teacher in the present experiment, however, observed that only the first and fourth of these questions were within the range of her students, and the fourth only with some sort of rephrasing. Ultimately, four questions were chosen (given in English to prevent any possible confusion) that were acceptable to the control teacher as well as to the experimental teacher. The questions focus on description and self-identification (*Figure 5*), though description and self-identification are far from the predominantly narrative focus of the mini-course. If anything, then, these questions were biased in favor of the control group, and the comparability of the two groups is striking.

Figure 4

Guten Tag! Wie geht's?
Wo ist die Bank?
Sie sind in einem Restaurant, und wollen essen. Was sagen Sie?
Haben Sie ein gross Haus, oder ein klein Haus? Ist Ihre Haus schön?

Figure 5

Hello. How are you?
Who are you? What are you like?
What do you like?
What do you like to do?
Describe your house.

The quantitative difference between the experimental and control groups appears in a more concrete dimension, discourse length. The number of words for each answer was totaled for each student (see *Figure 6*). A large difference was consistently seen in the individual students' answers and was not the result of a few loquacious students.

Figure 6

experimental	*control*	
52.8	*36.2*	*average total number of words per student*

The rapid writing test, with time limited to one and a half to two minutes, depending on the question, served as a useful and convenient test of speaking fluency. I have found it a fairly reliable, if indirect, reflection of student abilities in several years of using it in my Hebrew program. The biggest drawback seems to be that it bypasses certain kinds of shyness. Some students who freeze, either in front of a class or a microphone, are able to function in a rapid writing test. It is apparently somewhat easier than an oral test but is generally parallel, reflecting basic abilities to produce spontaneous discourse.

Of further interest is the grammatical portion of the test, although it could not be given to the control class. This portion consisted of a simple paragraph with grammatical endings left blank for students to fill in, providing two or three instances of various endings for present tense conjugation; gender and case (nominative or accusative for direct object; and dative for location). The control group had not covered enough of this grammar for the teacher to be willing to give it; in the experimental group, the averages were as shown in *Figure 7*. Although this is not a test of acquisition, results are close to what one might expect in natural acquisition. Indeed, transcripts of another class studying German revealed lesser but parallel success (conjugation, 86; gender, 57; case, 36), although the medium here was spontaneous speech rather than rapid writing, and testing was done after a shorter time (9 as opposed to 15 class hours). It is worth emphasizing that the mini-course, though not designed for grammatical purposes as such, was able to efficiently teach this much grammar; furthermore, the implication is that, to some extent, it feeds this grammar into the acquisition (rather than merely the learning) of the students.

Figure 7

conjugation	91%
gender	76
case	46

Phonetics/Orthography and Vocabulary

In developing the curriculum, I was finding a need for special asymmetric sequencing of phonetics/spelling, and lexicon. For example, German *ie* versus *ei* needed to be presented separately. The traditional approach was to provide a trick to distinguish them: pronounce the second vowel of the combination as in English; *ie* is pronounced like English long *e*, *ei* is pronounced like long *i*. This sort of solution was deemed unacceptable in the present framework because it does not provide immediate competence, but rather necessitates mid-sentence

calculation. The student must analyze in order to pronounce. An asymmetric solution to the problem of tone in Chinese may be an appropriate example for the Western reader (see Bar-Lev, 1991c.).

Even the lexicon soon showed a need for asymmetric sequencing because words could be easily confused if they were in any way phonologically or even just semantically similar, e.g., Greek *tragoudei, hhorevei* (sing, dance). I soon rejected the traditional presentation of the numbers 1 to 10, for example, because they seemed to create mid-sentence calculation, albeit of a special sort (e.g., "There are one, two three, four, five, six . . . yes, six books on the table"). Instead, I began presenting them in groups for maximum individuation of the different numbers: "1, 2, 5, 10, 100, 1000" in the context of buying; "3, 6, 9, 12" in the context of time of day.

This strategy is a straightforward solution to the general problem of confusion, although one that traditional teachers resist. Thus, one teacher using my mini-course in German, coming to *wer* (who) on page 24, felt constrained to direct the students' attention to page 45, to tell them "not to confuse *wer* with *wo* (where)" —another self-defeating instruction. Rebecca Valette (in a personal communication) reports a similar reaction among traditional teachers to her innovative approach (Valette & Valette, 1980) of separating *ser* "be (permanently)" and *estar* "be (temporarily)", which I also do in my Spanish mini-course.

Settling on the number of words to present at one time was also difficult. I have settled on five words or so at a time, in 15-minute lessons. In a college course, students can readily cover 80 to 100 or more words per week in a mini-course. This is no small achievement, since I am talking about a level of confidence at which students can be said, on average, to acquire more or less all of these words. Thus, timed vocabulary quiz scores regularly fall into the lower half of a bell curve, with the mean score at or near 100 percent; for example, on a recent Hebrew exam for the whole mini-course, grades out of a total of 20 averaged 17.3, with 7 students out of 19 scoring 20. In addition, students are able to speak spontaneously, using this vocabulary at a continuous pace. This pace continues through the 150 words of the mini-course, within fewer than 20 hours of class time. The vocabulary is generic; it may not include idioms for name or age, and it purposely excludes days of the week, months of the year, or other long lists. A typical vocabulary for the mini-course, by part of speech, for Hebrew, is given in *Figure 8*.

The grammar covered for any language is likely to include the most important persons in the present tense *(I, you-polite*, etc., but not *we* or *you-familiar)*; gender; and exposure to the plural and the infinitive, with up to three cases covered in case languages.

Figure 8

fish, meat, apricot, fruit, bread, milk, coffee, cake, flatbread, egg, wine, tea, orange, juice, soup, bottle, bill, newspaper, book, garment, chair, shekel, tree, letter (to someone), tablecloth, shirt, gift, map, watch/clock, medicine, gold, ticket, folder, garden, head, mouth

house, school, cafe, bank, market, hospital, street, library, station, mountain, America, Haifa, Russia, story, peace, war, rain, luck, music

doctor, teacher, friend, thief, Dad, Mom, soldier, son(s), daughter(s), king, pupil, prophet, man, woman, police, camel, elephant, pig, people

come, run, go up, go out, stand up, sing, cry, understand, drink, buy, (to) fast, see, bring out, put, hit, reside, cost, know, speak, eat, learn/study, sleep, write, receive, say, wear/put on, believe, like/love, can, want, work, help, it hurts, it's hot

good, big, little, good-looking, new, black, green, red, white, crazy, closed, American, English, Hebrew, Jewish, sad, satisfied, alive, obligated, strong, rich, dead, sick, fat

hello, welcome! good-bye, please, thanks, no, not, yes, okay, there isn't, there is, only, also, but, all, who?, you (to a male), you (to a female), I, he, she, they, we, to whom? to me, to you, what?, and, this is, from, of, with, without, how?, well, fast, in English, in Hebrew, where?, to, here, close, in, on, when?, at --- o'clock, Saturday, holiday, at night, during the day, the hour, how much?, a lot, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 18, 20, 30, 60, 70, 100, 1000

The typical student of the mini-course knows most of these words well enough to combine them freely in spontaneous speech, although students vary considerably in the imagination with which they do so. Volubility is a characteristic that all observers of SILL comment on: Students of the mini-courses are able to continue talking for four or more sentences—many can continue indefinitely—without hesitating, self-correcting, or the "humble-fumble" that often characterizes students even at more advanced levels.

Students of the German experiments were "ebullient," to recall the word used in Oxford et al (1990), and requested that a mini-course again be given to them before the second semester; several decided to major in German after these few weeks of study. Later, however, when they continued in *Deutsche heute*, a number of students informed me of their decision to quit the study of German.

Mini-course students of all ages and in all languages react enthusiastically; to paraphrase evaluations, they appreciate the control that the method gives them in class and in the language; they estimate it as many times more effective than other language courses they have had. Some find it intense because it demands individual speaking in groups and in front of the class from the first day, but they quickly get involved in the study steps and emerge empowered by the experience.

Students individually commend the method for giving them the "music" of the foreign language, or for making them feel (in the talkathons) as if they were in the foreign country and were suddenly able to speak in the language without stopping. Students like the way SILL keeps grammar in proportion to skill development and the way it allows them to speak rather than just to read or fill in worksheets. They support the way it puts them in charge of learning, so that they understand how to teach as well as learn a foreign language, even a new one. They often express a wish that other teachers would use the method.

Perhaps the most dramatic of the individual observed successes was my student "J," as observed by "A," a student of Japanese who had lived for two years in Japan. "A" estimated that "J" spoke as confidently as he himself had been speaking after one and one-half months of intensive study (10 hours a day): "J" had also studied for a month and a half, although only one hour a week. "A" did note that "J"s vocabulary was limited, but described his grammatical accuracy as "fine."

Other individual cases can serve to illustrate the effects of mini-courses on students who come to them from ordinary frameworks or later continue in such approaches. One student, "S," had been "kicked out," as she put it, of French and then Spanish in high school because of her supposed lack of ability. After an individual mini-course involving six half-hour meetings, "S" was speaking limited and imperfect Hebrew; however, a tape recording of her spontaneous speech was identified by a seminary professor as that of a third-year student. She later went into a traditional Hebrew course (audio-lingual with emphasis on conjugation) and earned a grade of B-minus.

Another student, "W," who had earned D- to C-range grades in high school Spanish, began as the most hesitant, self-correcting student I have seen. She studied Russian in a 6-week (18-hour) mini-course, unable at one point to finish a Russian sentence without cursing, but later reached the desired fluency and went on to ordinary Russian courses which the instructor (a certified ACTFL oral-proficiency interviewer) described as proficiency-oriented. "W" noted that she was more confident than other students through the three subsequent years of conventional study. In the Soviet Union for a summer program following her first year, she was the only student who spoke Russian voluntarily, and who sought occasions to use her Russian, making Russian friends, while other students in the program tried English in stores, and when forced to speak Russian, would self-correct too much to proceed effectively. In spite of her poor high school language grades, "W" earned A-range grades through two years of college Russian, but began sinking to B's in the third year conversation course, where she felt the teacher corrected her excessively because of her relative fluency.

Some among my colleagues seem to feel that my innovations are worthwhile for especially weak students, such as the learning disabled,

but that the average student doesn't need them. One proficiency-oriented teacher who observed my class noted that speech and writing were roughly at the level of his students at the end of the first semester or the beginning of the second, apparently not noticing that they had studied for only 14 class hours at the time, without benefit of a fluent teacher, or input, or a large class. His class, as I later had occasion to observe, was more fun than mine: Rather than students trying to think what they could say, they were given a list of funny questions with which to interview each other, and circulated around the class enjoyably filling them out. The level of language and content was higher in this class—but it originated from the teacher, not the students.

Mini-courses have been taught by others with minimal preparation and with similar results. Both new and experienced teachers have used the program in Spanish, German, Japanese, Chinese, and Hebrew, with little or no preparation and with dramatic results. A new teacher can be trained in a few hours; the best training seems to be studying a new language in the method.

This concludes the account of accomplishments and reactions to the SILL project, anecdotal as it must necessarily be. The SILL project seems to have adequately accomplished its goal of teaching speaking confidence. Moreover, it teaches enough fluency and grammar to recommend it as an initial unit, preceding ordinary study. A few weeks' investment, teaching words and structures that an introductory course must teach in any case, would seem to be a small price to pay to provide benefits that may not be available in ordinary frameworks, despite their claims as proficiency-oriented or communicative approaches and their goals of proficiency.

Realism

Skim Reading and Skim Listening

The SILL project, based by the nature of its lexical starting point on a non-traditional output orientation, was at variance with the natural approach and with most modern methods. While there is some research defending the importance of output (Chaudron, 1988; McLaughlin, 1987), none is as radical as SILL was from its inception, since by its nature SILL could not incorporate much input. The output approach arose from two impetuses: (a) the observation that children who are not forced to respond to their parents in a second language, but who grow up comprehending, do not develop much competence in the language; and (b) my own experience in acquiring Hebrew as an adult, in which input played a significant but limited role.

Receptive skills remained a problem for the SILL project, since it proclaims the importance of limiting vocabulary and grammar for more

successful study. But even in an ordinary input-based curriculum, the problem of receptive skills is not really solved. For example, Nation (1990) estimates that the vocabulary needed for actual comprehension of ordinary reading materials approximates 7,000 words (The figures are given for English, but certainly similar magnitudes are involved for other languages.) Can any learner, even in Spanish, acquire more than 2,000 words in a semester with sufficient spontaneity to comprehend ordinary written tests at a decent pace—not to mention oral discourse? Thus, even in Western European languages, a far slower pace must be accepted, unless the degree of acquisition is to be seriously compromised. The solution, to the extent that there can be any, must lie in the method used to achieve comprehension. Ordinary native-like comprehension cannot work for most students within these time limits. Students may acquire the ability to read with a dictionary without developing more realistic skills, unless different definitions of the receptive skills are developed for L2 learners.

The key to this problem, as it developed, was embedded in some earlier research developing the practical applications of reading theory. These included the innovations for teaching Hebrew and Arabic (bar-Lev, 1989), as well as other applications suitable for use in Russian and in languages with less difficult writing systems. The program now incorporates the teaching of skim reading and "skim listening" from the beginning, in a special sequence that begins with word-search puzzles like those sold at supermarket checkout counters, and culminates in a variety of comprehension and interaction activities.

The underlying assumption is that comprehensible input, at least in an LCT language, does not come naturally or easily. Teachers cannot ignore the grammatical and lexical features of texts given to elementary students. These texts cannot be authentic at early stages. The emphasis is on natural *skills* rather than natural texts. What is the advantage of using an authentic text if the methods used with it are artificial? Reading with a dictionary, with the teacher's help, with leading questions, or even with such pre-reading helps as vocabularies, are all approaches too artificial to be of any use in a natural environment; newspapers don't give out vocabularies in advance, nor do they always provide pictures or graphs.

In listening, discourses must not only contain a large proportion (20 to 50 percent or more) of unknown words; they must generally be presented only once orally, at a continuous, though not native, pace. In natural environments, after all, newscasts and other oral texts are not repeated. The goal of listening exercises should not be to teach students to hear "where words begin and end," as is often assumed, but rather to comprehend as well as they can without knowing all the words. Where listening exercises require students to note small differences from a written script, they should be limited to advanced students.

Not all classroom activities need to be authentic in the present sense. There is an important place for activities that increase vocabulary through guessing in context. Guessing in context, however, is a *learning* skill, as distinct from a using skill. The mini-courses now contain abundant authentic activities which build these skills, so that students will be prepared for real language use from the outset. In ordinary classrooms, intermediate or advanced students and their teachers may pay lip-service to guessing from context, but in actuality they often read or listen to texts so closely, aiming at such specific understanding, that it is impossible for them to read in an authentic way. They may persist in this approach even when the teacher does not require it.

Group work, too, is unrealistic in this context, since it tends to depend on the skills of the most advanced learners. Of course, group work should as much as possible be realistic in structure. Students should never be assigned the task of creating a response to a conversational cue, since this task constitutes an unnatural dissection of the reality. Videodisc technology, too, is unrealistic in this sense: Analyzing a conversation, even without being able to pull up the written version and observe it, is unforgiveably opposed to what one must actually do in a conversation. This is not to deny the possible uses of the technology, but only to emphasize their marginal role in developing authentic skills.

Language in Real Time

SILL is heretical in that it does not attempt to lower the affective filter; on the contrary, some first-day activities are calculated to produce tension of a specific kind, as in the "noisy restaurant" exercise, in which each student must shout out an order for food or drink, to the accompaniment of the class whispering nonsense syllables, with the teacher pacing in front of the class like a hurrying waiter. Forcing individual speaking is an unforgivable sin to some teaching theorists, and certainly it does not lower the affective filter. My view is that students must be capable of aggressive use of the language. What use is knowledge of a language if you cannot approach a group of native speakers engaged in conversation, interrupt, and ask a question, all because your skills are severely inhibited by such a demanding situation?

It might be argued that students can better approach such extreme challenges after they have acquired the language to a considerable degree in gentler environments. But my approach is that input itself is intimidating. By hearing a lot of the language before they produce, students learn how bad they sound before they speak. In SILL, students are taught from the beginning to feel comfortable with being a

funny-sounding foreigner. For this reason, SILL students exhibit confidence when speaking with native speakers after a few weeks or less. Although without illusions that they are speaking profoundly or even fluently, they are able to hold their own with native speakers on an interpersonal level, although not on a linguistic level.

One of the rules that SILL students learn is not to allow the native speaker to become a teacher; accordingly, they are not allowed to say, "Please repeat," "I don't understand," or even "What?" just as they are forbidden to self-correct or fumble—all of which can put them into an unnaturally dependent position with native speakers. When these forbidden communication strategies are disallowed, students, I find, speak quite naturally. I would classify these communication strategies instead as indicators of communicative breakdown. At advanced levels with graduates of mini-courses, I may be more flexible about allowing questions, but I have found it beneficial to enforce some version of the rules for advanced students who have come out of more standard curricula. These students may, in the name of accuracy, ask too many questions and give a poor impression of their abilities. As for the danger of reacting incorrectly, for example if someone shouts "Freeze or I'll waste you," the solution would seem to be to teach pragmatic awareness, as the mini-course does; no course can cover all possible meanings of all words.

The mini-courses encourage expression over and above communication as such, and thus encourage creativity. I have found that students have no trouble speaking sensibly when they know enough language to do so. But they do have trouble talking nonsense, and must be strongly encouraged to do so if they are to develop creativity. Creativity, proclaimed by Chomsky to be the hallmark of human speech, is not the exceptional creativity of a Shakespeare, but rather the ability of humans to insert lexical items into sentence patterns; hence his classic sentence, "Colorless green ideas sleep furiously." SILL does discourage such extreme nonsense; even science fiction such as "The table kisses the teacher" is likely to be interpreted as "The teacher kisses the table" in languages where word order plays a lesser role than in English.

In the mini-courses, creativity develops fairly linearly out of pattern practice. The three study steps of vocabulary quiz, pattern practice, and talkathon (talking for as long as possible, in a group or in front of the class) are adequate to develop the needed abilities, along with some special activities in review sections, for skim reading and grammar review.

As an output-oriented method, it is not surprising that SILL makes few linguistic demands on the teacher. It would seem to offer a solution to the traditional problem of teaching speaking abilities where the teacher is not fluent. There are also techniques by which ordinary

textbooks can be brought into better conformity with SILL, allowing development of speaking abilities with as little as two half-hour sessions a week. But even without special techniques, and even in a class in which the teacher does nothing to encourage spontaneous speaking, it would be a boon for students to enter regular study with elementary fluency of the type described here.

With the current push toward speaking abilities, I wonder whether any of the methods ordinarily proposed for teaching speaking can achieve the desired results, when they have to run the gauntlet of publishers, textbooks, programs, and teaching assistants. How can the 80s be characterized as the decade of oral proficiency, as suggested in Harlow and Caminero (1990), on the basis of 50 voluntarily responding departments in commonly taught languages stating that they examine orally an average of once a semester, for an average of 10 percent of the course grade? In the mini-courses, I have examined as many as 35 students in 75 minutes in up to seven languages, and 30 students in 50 minutes in a single language. I have even done a cursory oral exam for 70 students in 15 minutes. More important, since speaking is the predominant activity in the mini-courses, I have ample opportunity to observe students (and little need to do anything else for a good part of the class time).

The mini-courses aim at empowering the student individually; with group work, students become their own teachers. The SILL method is simple enough to learn quickly; the teacher's role can be mostly limited to providing pronunciations of words, monitoring to make sure that students keep on task, and nodding approvingly during talkathons. Correcting errors during spontaneous conversation does not improve accuracy; it merely undermines spontaneity (even if it improves accuracy in the resulting—less spontaneous—speech).

The goal of developing realistic skills quickly and as much as possible in "real time" seems a worthy one. It may seem radical in terms of current practice to impose this requirement on even a portion of class time, but how can we claim to teach the use of language if students are not capable of using it in real time? Where teachers are allowed to help students during exercises, or students are allowed to help each other, we cannot know how they will perform in real situations.

Similarly, the mini-courses avoid the topical consistency of the classroom, which primes students for specific schema sets and vocabulary sets in ways that ill prepare them for real life. Their generic focus is crucial: If students proceed from one situation to another, their abilities may never amount to more than the sum total of the situations they have studied.

Authenticity of skill guarantees that the material learned will be usable at earlier points than with ordinary curriculum. While the artificial skills imparted in ordinary curriculum are closer to authentic skills than the purely academic curriculum of grammar/translation methods, the allowability of instant replay, not to mention slow motion, in ordinary classrooms guarantees a certain unreality. This, combined with the topical organization of most curricula, prevents teachers from developing a clear picture of how their students will respond to "real" reality.

We often have strange misconceptions about what happens in our own classes. Many teachers, although communicatively oriented, unconsciously practice ventriloquism, finishing students' sentences and then crediting them to the student. I think we need to campaign for a more specific, more realistic concept than proficiency.

Appendix

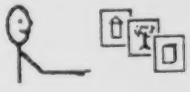
SAMPLE LESSONS
Hebrew

STUDY STEPS

- WORD QUIZ

Make sure that you can say each word instantly when cued by its English equivalent (or a picture, or the object itself).

Get someone to quiz you *until you are fluent.*



- PATTERN DRILL

Practice saying sentences.

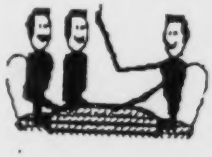
Talking is a sport: learn to do it *smoothly and confidently.*

Learn to talk without sputtering, hesitating, or backtracking.

Do pattern drill rhythmically; correct as necessary, but practice until rhythmic.

- TALKATHON

Talk for as long as possible without stopping or self-correcting.



BEST COPY AVAILABLE

LESSON 1
"please"

LEARNING TRICKS: match the MEANINGS with the **pronunciations**, and then learn the words.

dog FISH

APRICOT **mishmash**

cough at COFFEE

the ORANGE has **top-ooze**

low? NO!

PLEASE **back a shah**

WORDS

write translation here ↓	write translation here ↓
dahg **tapuz** **mishmish** **kafeh**	**ʿvaqashah** **lo**

SAMPLE SENTENCES

dag, ʿvakashah!

mishmish? lo.

MAKE UP SENTENCES OF YOUR OWN.

Use your new vocabulary to order whatever items you can.
Order very loud, or you won't be heard.

LESSON 2
"who?"

LEARNING TRICKS: match the MEANINGS with the **pronunciations**, and then learn the words.

WHO is **Mi**mi?

he COMES to the **bar** (no r)

a (male) TEACHER makes **more** effort

row, fetch the DOCTOR

peel an ELEPHANT

a PROPHET says "C'est **la vie**!" (l > n)

WORDS

moreh	**mi**
rofé	**bá**
naví	
pil	

SAMPLE SENTENCES

mi bá? moreh bá

rofé bá? lo, rofé lo bá.

All the words studied so far refer to males, for grammatical simplicity. You will learn how to refer to females below.

FREE YOUR IMAGINATION WITH ANIMAL STORIES

. . . such as you might tell a child. The goal is to increase your imagination for its own sake.

LESSON 3
actions

LEARNING TRICKS: match the MEANINGS with the **pronunciations**, and then learn the words.

charlotte SINGS

he RUNS and **rots**

WORDS

'sha^{h}r
ra^{h}ts

SAMPLE SENTENCES

moreh kam? a teacher is standing up?
lo, moreh rats.

MAKE UP SENTENCES OF YOUR OWN.
Practice saying sentences of your own until you can do so fluently.

REVIEW

WORD QUIZ. Translate each word; review until you can do so instantly.

fish, orange, apricot, coffee
who?, teacher, doctor, prophet, elephant
comes, runs, sings
please, no (not)

Make sure you can say a sentence instantly with each word.

Speak for as long as you can without stopping or hesitating.

LESSON 4
females

How many of the following words for females can you recognize? (You have learned them all, as used for males.)

WORDS

nevi'áh	**sharah, kamah, ratsah**
rof'áh	**yafah**
moráh	

SAMPLE SENTENCES

neviah sharah. rof'ah ba'ah. morah kamah.

MAKE UP SENTENCES OF YOUR OWN.

Practice saying sentences of your own until you can do so fluently.

Don't self-correct if you do make a mistake, as you will inevitably do every once in a while: Mistakes in endings don't affect meaning, and a Hebrew speaker will be able to understand you better even if you never use the right ending, rather than if you self-correct while speaking.

LESSON 5
"what?"

LEARNING TRICKS: match the MEANINGS with the **pronunciations** and learn the words.

he SEES a **row** (+ eh)

he BUYS a **cone** (+ eh)

WHAT is **Ma** doing?

WORDS

ro'eh	**mah**
koneh	

SAMPLE SENTENCES

navi ro'eh dag. morah ro'ah mishmish.

NOTE that only the verb ends in **a**^{h}, and only when the subject (the first noun in the sentence) is female. The object (the word after the verb) doesn't change, and the verb doesn't change for it.

MAKE UP SENTENCES OF YOUR OWN. Practice saying sentences of your own until you can do so fluently. Don't self-correct if you do make a mistake, as you will inevitably do every once in a while.

REVIEW

WORD QUIZ: Translate each word; review until you can do so instantly.

what?, fish, orange, apricot, coffee

who?, teacher-male, teacher-female, doctor-male, doctor-female, prophet, prophetess, elephant

comes, runs, sings, sees, buys

please, no (not)

Make sure you can say a sentence instantly with each word.

Speak for as long as you can without stopping.

References

Anonymous. (1990). Confessions of a textbook author. *Polylingua, 1*, 1.

bar-Lev, Z. (1989). Glyphs. *Bulletin of Hebrew Higher Education, 4*, 1.

bar-Lev, Z. (1991a). Pre-Hebrew. *Bulletin of Higher Hebrew Education, 4*, 4.

bar-Lev, Z. (1991b). Innovations for the 'other' teacher. *Shofar, 9*, 3.

bar-Lev, Z. (1991c). *1991, October*. Two innovations for teaching tones. *Journal of Chinese Language Teachers Association*,

bar-Lev, Z. (1991d). Two innovations for teaching Arabic. *Journal of Arabic Language Teachers Association, 24*.

bar-Lev, Z. (1992a, June). Skim-reading, skim-listening, and spontaneous conversation in elementary Arabic. Teaching Arabic in the 90s: Issues and Directions, Middlebury College.

bar-Lev, Z. (1992b, November). Skim-reading and skim-listening in elementary foreign language. ACTFL Annual Conference.

Blair, R. W. (1991). Innovative approaches. In Celce-Murcia (Ed.), *Teaching English as a second language* (pp 33-36). Rowley, MA: Newbury.

Brumfit, C. (1984). *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.

Celce-Murcia, M. (Ed.). (1991). *Teaching English as a second or foreign language (2nd ed.)*. Rowley, MA: Newbury.

Chaudron, C. (1988). *Second language classrooms*. Cambridge: Cambridge University Press.

Corder, S. (1967). The significance of learner errors. *International Review of Applied Linguistics, 5*.

Cushing, S, Taylor, H., Webster, C. & Schumann, J. (undated). A lexical method for language learning. UCLA.

García, C. (1989). Development of sociolinguistic skills: Sample activities, *Foreign Language Annals, 22*, 4.

Groesbeck, Col. W. (1988, December). Our burgeoning linguistics gap, *Army*.

Hammond, R. (1988). Accuracy vs. communicative competency: The acquisition of grammar in the second language classroom. *Hispania, 71*.

Harlow, L., & Caminero, R. (1990). Oral testing of beginning students at large universities: Is it worth the trouble? *Foreign Language Annals, 23*, 6.

Kennedy, D., & De Lorenzo, W. (1985). *Complete Guide to Exploratory Foreign Language Programs*. Lincolnwood, IL: National Textbook.

Krashen, S., & Terrell, T. (1983). *The natural approach*. San Francisco: Alemany.

McLaughlin, B. (1987). *Theories of second language learning*. London: Arnold.

Moeller, J., & Liedloff, H. (1984). *Deutsche heute*. Boston: Houghton Mifflin.

Nation, I. (1990). *Teaching and learning vocabulary*. Grand Rapids: Newbury.

Oxford, R., Crookall, A., Lavine, R., Nyikos, L., & Sutter, W. (1990). Strategy training for language learners: Six situational case studies and a training model. *Foreign Language Annals, 23*, 3.

Pritsky, E. (1987). Russian and the pre-literate learner: On the development of a methodology, *Foreign Language Annals, 22*, 3.

Porter, L. Using the ACTFL proficiency guidelines to achieve goals of exploratory language courses, *Foreign Language Annals, 20*, 4.

Reinert, H. (1970). Student attitudes toward foreign language—no sale!, *Modern Language Journal, 54*.

Rutherford, W. *Second language grammar: Learning and teaching*. London: Longman.

Schumann, J. (1986). Second language acquisition: The pidginization hypothesis, *Language Learning, 26*.

Selinker, L. Interlanguage. (1972). *International Review of Applied Linguistics, 10*.

Simon, Paul. (1988). *The tongue-tied American*. New York: Continuum.

Terrell, T., Andrade, M., Egasse, J., & Muñoz, E. (1990). *Dos Mundos. NY: McGraw-Hill*.

Valette, J-P., & Valette, R. (1980). *Spanish for mastery: I*. Lexington, MA: Heath.

VanPatten, W. (1992). Structured input and structured output: Explicit grammar instruction for the communicative classroom. ACTFL Annual Convention, Chicago.

Author

ZEV BAR-LEV, Professor, Department of Linguistics and Oriental Languages, San Diego State University, San Diego, CA 92182. Specializations: applied linguistics, (TESOL, foreign language teaching), Hebrew.

Calendar of Events *

1993

Japan Association of Language Teachers, 8-11 October, Tokyo. *Information:* JALT, Shamboru Dai 2 Kawasaki #305, 1-3-17 Kaizuka, Kawasaki-ku, Kawasaki 210, Japan; +81-44-245-9753, fax +81-44-245-9754.

Mountain Interstate Foreign Language Conference, 14-16 October, Clemson. *Information:* Sixto E. Torres, Dept. of Languages, 201 Strode Tower, Clemson University, Clemson, SC 29634-1515; (803) 656-3393, fax (803) 656-0258.

Illinois Joint Foreign Language Conference, 21-24 October, Peoria. *Information:* Susan Leibowitz, Glenbrook South High School, Glenview, IL 60025; (312) 549-7517.

Classical Association of the Atlantic States, 22-23 October, Annapolis. *Information:* Mary E. Moser, Department of Classical Studies, Dickinson College, Carlisle, PA 17013; (717) 245-1685.

Foreign Language Association of North Carolina, 28-30 October, Greensboro. *Information:* Wayne Figart, 204 N. 16th ST., Wilmington, NC 28401; (919) 763-4009.

Massachusetts Foreign Language Association, 29-30 October, Sturbridge. *Information:* Georg Steinmeyer, Black Mt. Rd. RFD #1, Box 549, Brattleboro, VT 05301, fax (802) 257-1855.

Youngstown Foreign Language Conference, 29-30 October, Youngstown. *Information:* Department of Foreign Languages, Youngstown State University, Youngstown, OH 44555; (216) 742-3461, fax (216) 742-1419.

Wisconsin Association of Foreign Language Teachers, 5-6 November, Appleton. *Information:* William W. Kean, Suring Public Schools, PO Box 158, Suring, WI 54174.

Northern New England TESOL, 13 November, Manchester. *Information:* Don Bouchard, University of Southern Maine, College of Education, Department of Literacy Education, Bailey Hall, Gorham, ME 04038; (207) 780-5069.

* Courtesy of *The Modern Language Journal* *(University of Wisconsin Press)*

American Association for the Advancement of Slavic Studies, 19-22 November, Honolulu. *Information:* AAASS, Stanford University, 128 Encina Commons, Stanford, CA 94305-6029; (415) 723-9668, fax (415) 725-7737.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 20-22 November, San Antonio. *Information:* ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, fax (914) 963-1275.

Federación Nacional de Profesores de Inglés de Universidades Y Politécnicas, 22 November, Guayaquil. *Information:* Fausto Saltos, CELEX-ESPOL, Apartado 6117, Guayaquil, Ecuador, 59-34-32-1925, fax 59-34-31-3287.

Modern Language Association of America, 27-30 December, Toronto. *Information:* Modern Language Assocation of America, 10 Astor Place, New York, NY 10003-6981.

1994

American Association of Applied Linguistics, 5-8 March, Atlanta. *Information:* AAAL Membership Office, PO Box 24083, Oklahoma City, OK 73124; tel (405) 843-5113, fax (405) 843-4863, e-mail (jmay@rex.chb.uokhsc.edu).

Second/Foreign Language Acquisition by Children, 10-12 March, Tulsa. *Information:* Rosemarie Benya, PO Box 2053, Ada, OK 74821; (405) 332-8000.

Computer Assisted Learning Instruction Consortium, 14-18 March, Flagstaff. *Information:* CALICO, Duke University Box 90267, Durham, NC 27708-0267; (919) 681-6455, fax (919) 681-6485, e-mail (CALICO@DukeMVS.ac.duke.edu).

Ohio Foreign Language Association, 17-19 March, Dayton. *Information:* Lori Winne, 6238 Summit Street, Erie, MI 48133.

Fédération Internationale des Professeurs de Langues Vivantes, 28 March - 1 April, Hamburg. *Information:* FIPLV Head Office, Seestrasse 247, CH-8038 Zürich, Switzerland.

Southwest Conference on Language Teaching with **Utah Foreign Language Association**, 31 March - 3 April, Salt Lake City. *Information:* Joann K. Pompa, Mountain Pointe High School, 4201 East Knox Road, Phoenix, AZ 85044; (602) 759-8449.

Renaissance Society of America, 7-9 April, Texas. *Information:* Craig Kallendorf, English Department, Texas A & M University, College Station, TX 77843.

Northeast Conference on the Teaching of Foreign Languages, 7-10 April, New York. *Information:* Northeast Conference, St. Michael's College, Dupont Hall, 29 Ethan Allen Avenue, Colchester, VT 05439.

Conseil International d'Etudes Francophones, 9-16 April, Québec City. *Information:* Maurice Gagnon, French Department, Montclair State University, Upper Montclair, NJ 07043; (201) 893-5143; fax (212) 666-3715.

Central States Conference on the Teaching of Foreign Languages with **Missouri Foreign Language Association**, 21-24 April, Kansas City. *Information:* Jody Thrush, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; (608) 246-6573, fax (608) 246-6880.

California Foreign Language Teachers Association with **Foreign Language Association of Northern California**, 22-24 April, San Francisco. *Information:* Sylvia Jones, 62 N. Arroyo Boulevard, Pasadena, CA (818) 356-0112.

Louisiana Foreign Language Teachers' Association, 22-23 April, Lafayette. *Information:* Jeannette Hicks, Tr. 6, Box 190, DeRidder, LA 70634.

Classical Association of the Atlantic States, 29-30 April, Carlisle. *Information:* Mary Moser, Department of Classics, Dickinson College, Carlisle, PA 17013; (717) 245-1683.

Pacific Northwest Council on Foreign Languages with **Montana Association of Language Teachers**, 12-14 May, Missoula. *Information:* Ray Verzasconi, Foreign Languages and Literatures, Oregon State University, Corvallis, OR 97331-4603; (503) 737-2146, e-mail (verzascr@ccmail.orst.edu).

Society for the History of Authorship, Reading and Publishing, 14-16 July, Washington. *Information:* John Y. Cole, Center for the Book, Library of Congress, Washington, DC 20504; tel. (202) 707-5221.

Interantional Assocation for Southeast European Studies, September date to be announced, Greece. *Information:* AIESEE, Department of Slavic Languages, CB #3165, Dey Hall, University of North Carolina, Chapel Hill, NC 27599-3165.

Foreign Language Association of North Carolina, 13-15 October, Greensboro. *Information:* Wayne Figart, 204 N. 16th St., Wilmington,NC 28401; (919) 763-4009.

Wisconsin Association of Foreign Language Teachers, 4-5 November, Appleton. *Information:* William W. Kean, Suring Public Schools, PO Box 158, Suring, WI 54174.

American Association for the Advancement of Slavic Studies, 17-02 November, Philadelphia. *Information:* AAASS, Stanford Univerisity, 128 Encina Commons, Stanford, CA 94305-6029.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 18-20 November, Atlanta. *Information:* ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, fax (914) 963-1275.

Modern Language Association of America, 27-30 December, San Diego. *Information:* Modern Language Assocation of America, 10 Astor Place, New York, NY 10003-6981.

1995

Central States Conference on the Teaching of Foreign Languages, with **Pacific Northwest Conference on Foreign Languages**, **Southwest Conference on Language Teaching**, and **Colorado Congress of Foreign Language Teachers**, 30 March - 2 April, Denver. *Information:* Jody Thrush, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; (608) 246-6573, fax (608) 246-6880.

Northeast Conference on the Teaching of Foreign Languages, 6-9 April, New York. *Information:* Northeast Conference, St. Michael's College, Dupont Hall, 29 Ethan Allen Avenue, Colchester, VT 05439.

American Association of Teachers of German with **Internationaler Deutschlehrerverband**, 4-7 August, Stanford Univerisity. *Information:* AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, fax (609) 795-9398.

Foreign Language Association of North Carolina, 12-14 October, Greensboro. *Information:* Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; (919) 763-4009.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 18-20 November, Anaheim. *Information:* ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, fax (914) 963-1275.

Modern Language Association of America, 27-30 December, Chicago. *Information:* Modern Language Assocation of America, 10 Astor Place, New York, NY 10003-6981.

1996

Northeast Conference on the Teaching of Foreign Languages, 18-21 April, New York. *Information:* Northeast Conference, St. Michael's College, Dupont Hall, 29 Ethan Allen Avenue, Colchester, VT 05439.

Foreign Language Association of North Carolina, 24-26 October, Greensboro. *Information:* Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; (919) 763-4009.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 22-24 November, Philadelphia. *Information:* ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, fax (914) 963-1275.

Modern Language Association of America, 27-30 December, location to be announced. *Information:* Modern Language Assocation of America, 10 Astor Place, New York, NY 10003-6981.

1997

Northeast Conference on the Teaching of Foreign Languages, 3-6 April, New York. *Information:* Northeast Conference, St. Michael's College, Dupont Hall, 29 Ethan Allen Avenue, Colchester, VT 05439.

Foreign Language Association of North Carolina, 10-12 October, Greensboro. *Information:* Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; (919) 763-4009.

Modern Language Association of America, 27-30 December, location to be announced. *Information:* Modern Language Assocation of America, 10 Astor Place, New York, NY 10003-6981.

Attention, associations: As soon as your conference dates and locations are booked, please send the information. We will post it in the *Calendar of Events* section of *Applied Language Learning*.

537

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editors encourage the submission of manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006.

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating:

full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

Specifications for Floppy Disks

Manuscripts should be typed on one side only on 8½ x 11 inch paper, double-spaced, with ample margins. Where feasible, manuscripts are preferred on disk (DOS or XEROX Viewpoint).

Manuscripts produced on one of the DOS-based or Macintosh systems should be entered on the disk as an *ASCII* file if possible. Preferred is the 5¼-inch size formatted as *MS-DOS* file at double density.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____
2. Saved as an ASCII file? YES_____NO _____
3. Disk is formatted as: double: _____high density_____
4. Remarks:

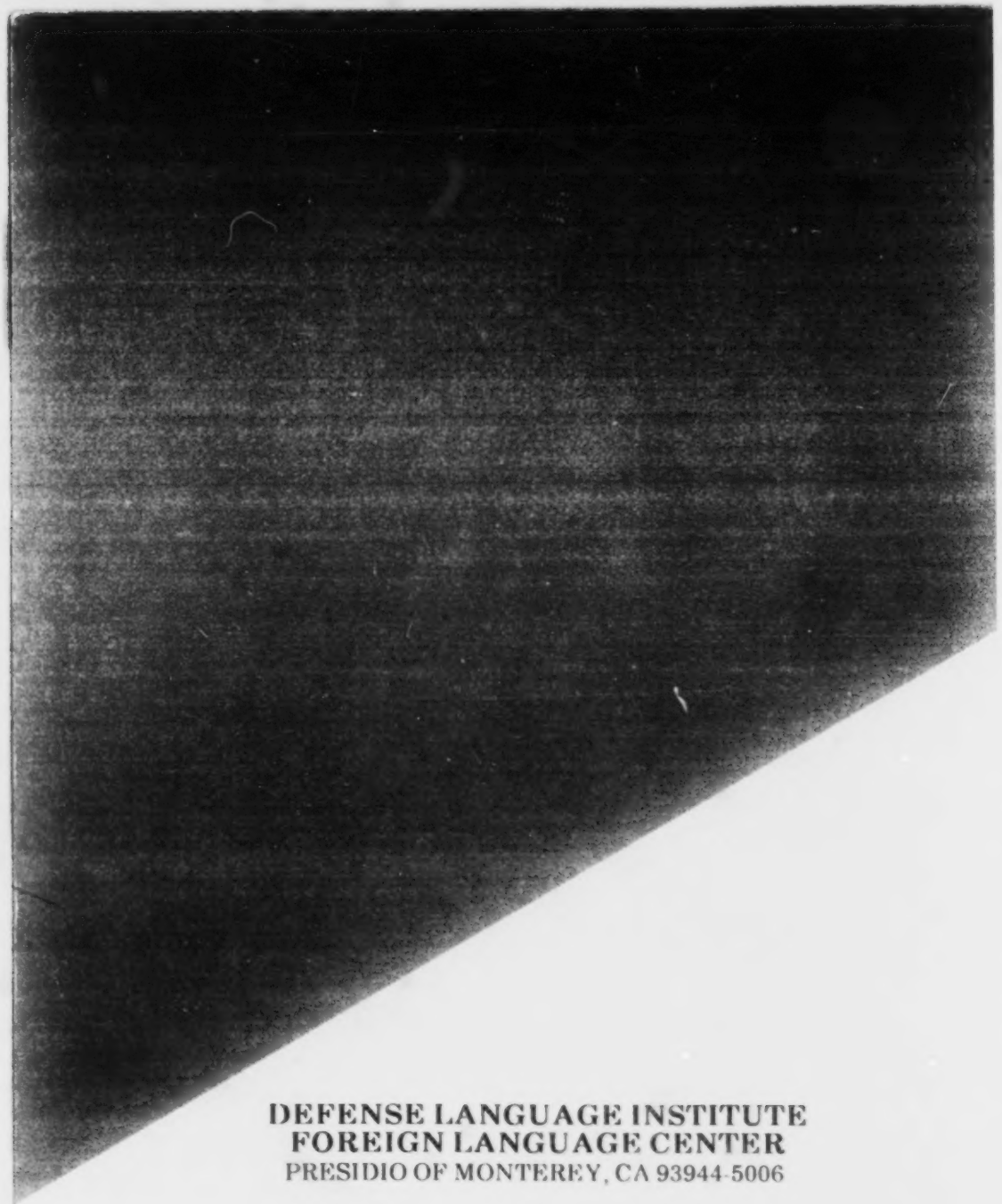
Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

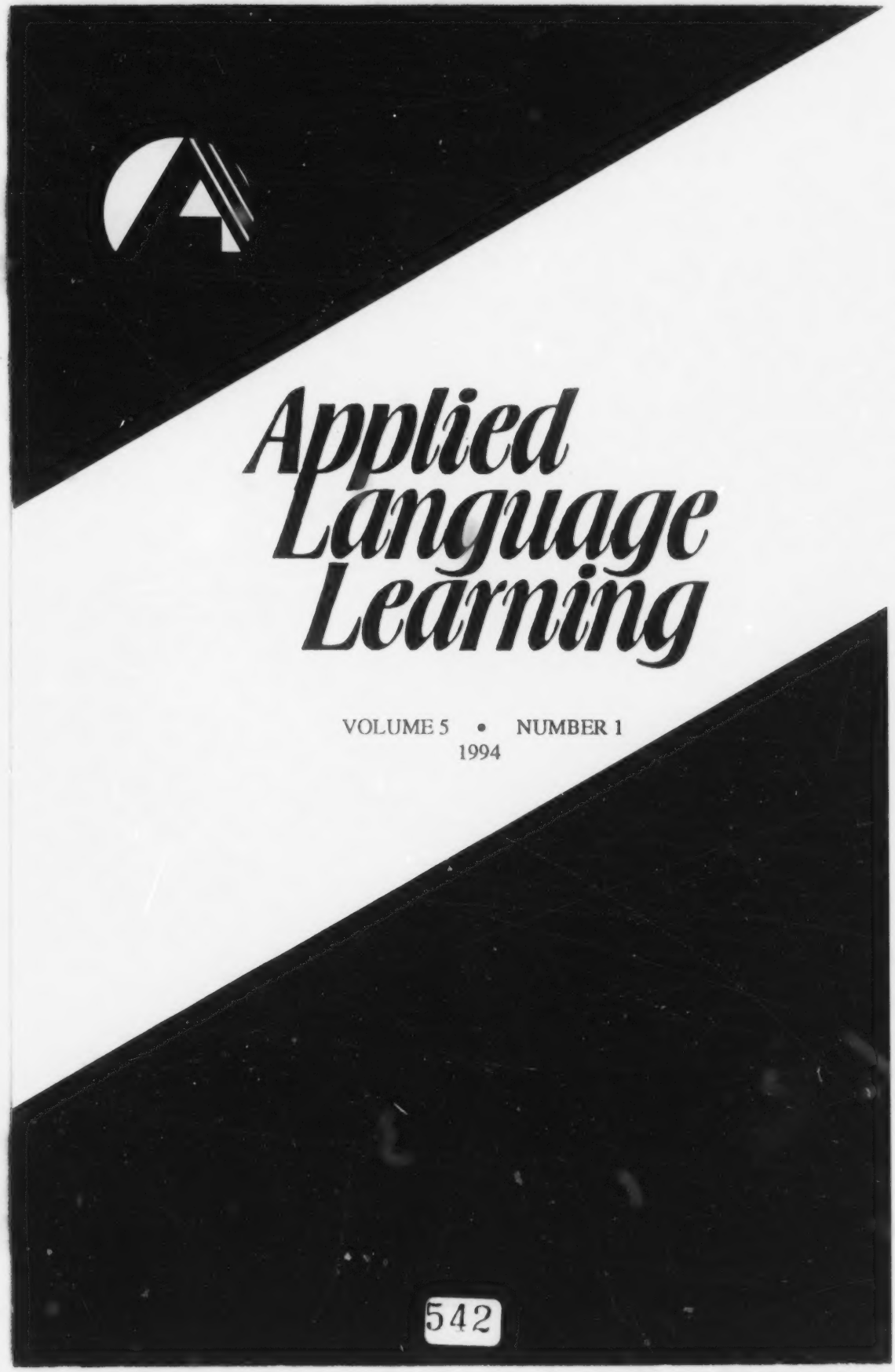
PIN:070788-UN-000



DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

541

Professional Bulletin 65-93-1 & 2
United States Army
Approved for public release. Distribution is unlimited.



Applied Language Learning

VOLUME 5 • NUMBER 1
1994

542

Applied Language Learning

Volume 5 ✣ Number 1

1994

Defense Language Institute
Foreign Language Center

543

APPLIED LANGUAGE LEARNING

PB 65 - 94 - 1

The mission of Professional Bulletin 65-94, *Applied Language Learning* (US ISSN 1041-679), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semi-annually by the Defense Language Institute, Foreign Language Center, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

Official:

MILTON H. HAMILTON
Administrative Assistant to the Secretary of the Army

GORDON R. SULLIVAN
General, United States Army
Chief of Staff

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to DLIFLC.

Applied Language Learning is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Book rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster — Send change-of-address information to: *Applied Language Learning,* Office of the Provost, Defense Language Institute, Foreign Language Center, Presidio of Monterey, CA 93944-5006. Telephone: (408) 647-5638 or DSN: 878-5638.

APPLIED LANGUAGE LEARNING

PB 65 - 94 - 1

1994 • Volume 5 • Number 1

Lidia Woytak, Editor

Editorial Board

Kathleen M. Bailey	Professor, MIIS
David P. Benseler	Professor, Case Western Reserve University
Christine Campbell	Project Director, Testing, DLI
John L. D. Clark	Dean, Program Evaluation, Research, and Testing, DLI
Ray T. Clifford	Provost, DLI
Martha Herzog	Associate Provost, DLI
Stephen D. Krashen	Professor, University of Southern California
John A. Lett, Jr.	Director, Research, DLI
June K. Phillips	Dean, College of Arts and Humanities, Weber State University
Gary D. Walter	Librarian, DLI

DLI Staff

Editorial Coordinator, Christa Rutsche
Production Editor, Jack Boyd
Associate Editor, Mike McGuire
Distribution, Roderick Wilson
Graphics, Ronald W. Williams
Layout Consultants, Annemarie Frazier and Bernhard Peters

545

APPLIED LANGUAGE LEARNING

PB 65 - 94 -1

1994 • Volume 5 • Number 1

CONTENTS

1 Factors in the Incidental Acquisition of Second Language Vocabulary from Oral Input: A Review Essay
Rod Ellis

33 Intake Factors and Intake Processes in Adult Language Learning
B. Kumaravadivelu

73 Pragmatics of Interaction: Expressing Thanks in a Second Language
Eli Hinkel

93 Deviant English and the Para-Poetic
Clark Lunberry

115 Calendar of Events

123 Information for Contributors

546

Applied Language Learning
1994, Vol. 5, No. 1, pp. 1-32

Factors in the Incidental Acquisition of Second Language Vocabulary from Oral Input: A Review Essay

Rod Ellis
Temple University

Although there is a growing volume of work on second language (L2) vocabulary acquisition, much of it has focused on intentional vocabulary learning or on incidental vocabulary learning from written texts. To date, little attention has been paid to how L2 learners acquire vocabulary incidentally from oral input. The main purpose of this article is to examine the factors that potentially affect incidental L2 vocabulary acquisition from oral input. Four sets of factors are considered: the intrinsic properties of L2 words; the nature of the input; the role of interaction; and individual learner factors. In addition, this article considers the nature of L2 vocabulary acquisition, and in particular examines differences in the comprehension and acquisition of L2 words.

Introduction

Second language acquisition research makes a traditional distinction between *incidental* and *intentional* acquisition. This distinction is manifest in a variety of terms, in particular *acquisition* versus *learning* (Krashen, 1981) and *implicit* versus *explicit* learning (Bialystok, 1978; Ellis, 1990). On the one hand, learners are credited with the ability to pick up L2 items and rules while their attention is focused primarily on trying to communicate (i.e., understanding and conveying messages). On the other hand, many learners also have the ability to focus their primary attention on the language code itself in order to try to learn new L2 items and rules. The two types of learning cannot be distinguished solely in terms of attention, though, for as Schmidt (1990) has convincingly argued, incidental learning necessarily involves a degree of consciousness when learners notice new items and rules in the input. The distinction rests, somewhat uncomfortably, on a secondary distinction between *focal* and *peripheral* attention: whereas intentional learning requires focal attention to be placed on linguistic form, incidental learning requires

© 1994, Rod Ellis

547

focal attention to be placed on meaning (i.e., message content), but allows peripheral attention to be directed at form.[1]

The distinction between incidental and intentional learning is of particular significance to the acquisition of vocabulary, and there is a considerable literature on the relative effectiveness of the two types of learning. In general, it appears that intentional learning that utilizes mnemonic strategies—such as the keyword method—results in better recall of word meaning than incidental learning (see Pressley, Levin, & McDaniel, 1987; Hulstijn, 1992).[2] However, it is also generally recognized, although not universally (see, for example, Goulden, Nation, & Read, 1990), that learners, even the most conscientious, can never acquire a native-like vocabulary through intentional learning. The argument is that no matter how much effort they put into deliberately attending to new words they encounter and no matter how effective their memorization or practice strategies are, they simply do not have the time to learn that amount of vocabulary intentionally. Therefore, they must learn a large portion of it incidentally. As Sternberg (1987, 1989) uncompromisingly states, "Most vocabulary is learned from context."

How does this incidental learning take place? The general assumption (see Nagy & Herman, 1987; Sternberg, 1987; Nation, 1990; Stoller & Grabe, 1993) is that it involves the utilization of various cues that enable learners to infer the meanings of new items from context. But there is a problem here, for as Coady (1993) points out, no study has yet been designed that shows how contextual learning really occurs.[3] Furthermore, there is a logical problem, again succinctly described by Coady:

> It is striking that the very redundancy or richness of information in a given context which, on the one hand, enables a reader to successfully guess an unknown word also predicts, on the other hand, that the same reader is less likely to learn the word-form because he or she was able to comprehend the text without needing to know it (p. 18).

To put it another way, if learners can successfully decode meaning without attending to the form of a new item, they may not notice the item and, if they do not notice it, they will not learn it. Inferencing from context, therefore, can obviously aid understanding, but may do nothing for learning.[4]

This problem, however, may reflect the preoccupation of researchers with *written* text as a source of input for vocabulary acquisition. Note that Coady refers to the *reader* in the quotation above. In the case of written text, the only cues available to learners for inferencing the meanings of new words are those contained in the text itself (although there may also be extralinguistic cues in the form of

pictures and diagrams). Often enough these cues are inadequate (see Stein, 1993). As Hulstijn (1992) has shown, many of the inferences that learners make from written context are incorrect. Oral text, particularly that which derives from face-to-face interaction, is different, however. In addition to the cues provided by the verbal text itself, there are other cues available from the intonation and gestures used by the speaker, and from the situational context. Furthermore, as L2 acquisition researchers like Long (1981) have emphasized, interaction affords opportunities for learners to indicate their non-comprehension and through the negotiation of meaning to receive input that is more finely tuned to their abilities to comprehend. What this often means, where difficult words are concerned, is that learners are provided with synonyms or definitions. Nagy and Herman (1987) are surely right when they claim, "written context will, therefore, generally not be as rich or helpful as oral context in providing information about the meanings of new words" (p. 24).

It seems reasonable to suppose that oral input may well constitute an effective source of data for incidental vocabulary learning. Indeed, the success which children have in developing a substantial vocabulary in their first language, before they come to school, testifies to this.

Given the importance of oral input for incidental vocabulary acquisition in many learning contexts, together with its potential to facilitate vocabulary acquisition, it is surprising that so little attention has been paid to it in L2 vocabulary acquisition research.[5] Whereas there are whole books given over to the relationship between reading and vocabulary (McKeown & Curtis, 1987; Huckin, Haynes, & Coady, 1993), and other books that address a variety of issues relating to L2 vocabulary (Nation, 1990; Arnaud & Bejoint, 1992), it is difficult to locate a single article on the relationship between oral input and L2 vocabulary development. An exception to this generalization can be found in the work that has addressed the effects of listening to stories on vocabulary acquisition (Wells, 1986; Eller, Pappas, & Brown, 1988; Elley, 1989; Feitelson, Goldstein, Iraqi, & Share, 1993). However, most of this research has concerned L1 vocabulary acquisition. Also, it has focused on what might be called *spoken written input* rather than naturally occurring oral input.

The purpose of this paper, therefore, is to make a start in examining how incidental L2 vocabulary acquisition can take place from oral input. To this end, it addresses this question: What factors govern the learnability of L2 lexical items in oral input? The aim is to provide an initial framework that can serve as a basis for empirical research. Before embarking on this framework, however, it is necessary to give some consideration to the dependent variable—vocabulary acquisition.

The Nature of the Learning Task—Vocabulary Acquisition

There is a quantitative and a qualitative dimension to vocabulary acquisition. On the one hand we can ask, "How many words do learners know?" while on the other we can inquire, "What do the learners know about the words they know?" Curtis (1987) refers to this important distinction as the *breadth* and *depth* of a person's lexicon. The focus of much vocabulary research has been on breadth, possibly because this is easier to measure. Arguably, however, it is more important to investigate how learners' knowledge of words they already partly know gradually deepens.

As Nation (1990) points out, there is considerable variation in the estimates of the breadth of native speakers' vocabulary. This reflects variation in the definition of a word (e.g., whether *govern, governing, governed, governor and government* are considered five different words or the same word), and differences in the way vocabulary is measured (e.g., whether recognition or recall of a word is required). Goulden, Nation, and Read (1990) suggest that the average educated native speaker has a vocabulary of about 17,000 base words, a figure that suggests a rate of acquisition of between two and three a day. A frequently cited figure, however, is Nagy and Herman's (1987) 40,000 for a senior high school student, which suggests students must add around 3,000 new words each year to their lexicon. Nagy et al. (1985) estimate that the rate of vocabulary acquisition necessary to achieve this figure is around .25 words per minute. Of course, there is considerable variation from one person to another.

Variation in the breadth of L2 learners' lexicons can be expected to be even greater, a reflection not just of the number of years they have been learning the L2, but also the extent and nature of their exposure to it and their purposes for learning it. Learners who need to develop cognitive academic language proficiency (Cummins, 1983) in order to study through the medium of the L2 at university level are likely to manifest greater breadth than learners who need only to develop the basic interpersonal skills to participate in day-to-day interactions.

A number of studies have investigated the size of L2 learners' lexicons (Barnard, 1961; Quinn, 1968; Jamieson, 1976 (all cited in Nation, 1990); Yoshida, 1978; Palmberg, 1987; Wode et al., 1992). More interesting, perhaps, than the absolute size of learners' vocabularies, is the rate at which they increase under specific learning conditions. Yoshida's study found that a 3-year-old beginner learner exposed to English for 2 to 3 hours a day at nursery school had acquired productive use of 260-300 words over 7 months. Wode et al. (1992) analyzed the productive vocabulary of German children acquiring English naturally in the U.S. and noted that the rate of acquisition tended to peak quite early

and then gradually decrease. This contrasts with L1 acquisition, where growth is slow up to 50 words and then accelerates. Barnard (1961, cited in Nation, 1990) found that school children in India learned between 1,000 and 2,000 words over a 5-year period. Palmberg (1987) found evidence of a steady increase in Swedish children's vocabulary as a result of their English studies at school (two 45-minute lessons per week), with the number of words recalled almost doubling after a 4-1/2 month period.

As we have already noted, the depth or quality of L2 learners' vocabulary knowledge is arguably of greater importance than its breadth. A learner may "know" 2,000 words, but this knowledge may be restricted to the recognition of these words in specific contexts. For example, a learner may understand the reference of *ruler* in one context:

I drew a line with my ruler.

but be incapable of recognizing the same word in a different context:

My ruler gives centimeters as well as inches.

and substitute another, incorrect, word in production:

Do you have a scale I can borrow?

Another learner may "know" only 1,000 words but may be able to recognize each word in a range of contexts and also use them accurately in production. As Twadell (1973) notes, "We may 'know' a very large number of words with various degrees of vagueness—words which are in a twilight zone between the darkness of entire unfamiliarity and the brightness of complete familiarity." But how can we characterize the depth of a learner's vocabulary knowledge more precisely?

A useful distinction is the one made between *potential* and *real* vocabulary (Palmberg, 1987). Potential vocabulary refers to L2 words that are not yet part of the learner's lexicon but can nevertheless be recognized when they are encountered because, for example, they are cognate with L1 words. Thus, a Japanese learner of L2 English may not know *ice cream* but is able to recognize it because it is close to Japanese *aisukurimu*. Real vocabulary consists of those words that have actually been "learned," to some extent.

It is traditional to distinguish receptive and productive (or expressive) knowledge. Receptive vocabulary consists of those words that a learner can recognize but may or may not be able to use; productive vocabulary consists of those words the learner can actually use in speech or writing. It is claimed that this distinction reflects a continuum rather than a dichotomy (Crow & Quigley, 1985; Palmberg, 1987; Coady, 1993). The assumption is that word knowledge is initially receptive and then, over time, through frequent encounters and practice, becomes productive. This need not be the case, however. It is perfectly

possible to envisage a learner having productive knowledge of a word without receptive knowledge. For example, when speaking Spanish, I can say *cosmopolita* (thus making my potential knowledge real by foreignizing a cognate English word), and yet be unable to recognize this item in oral input because I have not yet formed a phonological representation of the Spanish word.

A more useful way of representing depth of word knowledge involves the distinction between *knowledge* and *control* (Bialystok & Sharwood Smith, 1985). The application of this distinction to lexis is shown in *Figure 1*. *Knowledge* of a word requires the learner to:

- form a phonological/ graphological representation of a word
- understand its denotive reference
- understand its connotative meaning(s)
- know in what grammatical patterns it can be used
- discover with what words it collocates, and
- know with what other words it is typically associated

(see Richards, 1976, for a full account of these aspects of word knowledge). *Control* concerns the ability of the learner to access a word in reception and production.

Clearly, there can be degrees of knowledge and degrees of control. For example, Dale (1965) identified the following degrees of understanding, all of which would seem to apply to knowledge of a word's denotive meaning:

Stage 1: *I never saw it.*

Stage 2: *I've heard of it, but I don't know what it means.*

Stage 3: *I recognize it in context—it has something to do with . . .*

Stage 4: *I know it.*

Drum (1983) also distinguishes degrees of word-understanding on the basis of a person's ability to produce different kinds of definition. For example, a person may only be able to define a word by using it in a word or phrase, reflecting only a shallow knowledge of its reference (e.g., *the icons in the church*). Alternatively, the person might be able to define the word by giving an accurate paraphrase of its meaning (e.g., *an icon is a sacred representation of a holy person*). We can also envisage that the other aspects of word knowledge will vary in depth. For example, Meara's (1982) work suggests that the types of associations found in L2 learners' lexicons differ from the types found in native speakers; clang associations such as *clog/dog* are much more prominent in the former.

Irrespective of what learners know about a word, they are likely to vary in their ability to access their knowledge in accordance with the conditions of use. As we have already noted, they may be able to use a word receptively but not productively, or vice versa. They may be able to access a word (for reception or production) when given time to do so, but be unable to do so instantly. We are all familiar with the tip-of-the-tongue phenomenon.

FIGURE 1
A Model of Depth of Vocabulary Knowledge

KNOWLEDGE

CONTROL	form	denotive reference	connotative meaning	grammatical patterns	collocations	associations
reception						
production						

Clearly, depth of word knowledge is a complex issue, and the above discussion has done no more than skim the surface. It will not be possible to further our understanding until we heed Meara's (1984, 1992) call and develop adequate models of how a learner's mental lexicon is constructed by considering it as a whole—as "a large-scale structure"—rather than focusing on a small number of individual words. We will not proceed very far in our study of how incidental vocabulary learning from oral (or written) context takes place unless we have a clear understanding of what we mean by "vocabulary learning" and arguably we are still a long way from this.

Factors Affecting Incidental Vocabulary Learning

We can hypothesize that a number of factors will influence incidental vocabulary learning. Inevitably, given the paucity of research into acquisition from oral input, this account will have to draw on research involving written texts. It should, therefore, be seen as a programmatic statement rather than a survey. The factors are grouped under four headings: *intrinsic word properties*, *input factors*, *interactional factors*, and *learner factors*.

Intrinsic Word Properties

It seems very likely that some words will be more difficult to learn from oral input than others, irrespective of the richness of contextual

clues available to the learners. We will consider five intrinsic word properties which appear to influence acquisition.

1. Part of speech

There is some evidence to suggest that learners learn nouns more readily than other parts of speech, at least in the early stages. Yoshida (1978) found that his subject, a 3-1/2 year-old Japanese boy, learned mainly nouns initially—60 percent of his total words after 4-1/2 months were nouns. In contrast, he learned few verbs. Felix (1975) also found that his subjects (child learners of L2 German) were slow to learn verbs. Nation (1990) cites a study by Rodgers (1969) that found nouns were the easiest to learn, with adjectives next. Verbs and adverbs were the most difficult. Elley (1989), in a study investigating the effects of listening to stories read aloud on the acquisition of new words by 7- and 8-year-old native speakers of English, found a tendency for the children to show higher gains on nouns than on adjectives and verbs.

There are a number of explanations for these findings. One is, as Nation suggests, that the meanings of nouns can be guessed from context more easily than the meanings of verbs. Another very likely explanation is that learners initially concentrate on nouns because they are more useful for decoding and encoding messages. As Ellis (1984) has pointed out, early L2 utterances are frequently propositionally reduced because learners have limited processing capacity. Verbs are omitted more frequently than nouns because they are not so important for getting a message across. They may be attended to less in the input for the same reason. Nouns also regularly occur in sentence final position, whereas verbs occur less frequently in this position. There may be some kind of operating principle (Slobin, 1973)—for example, "Pay attention to the ends of sentences"—that accounts for why nouns are given precedence by beginner learners. Finally, N. Ellis and Beaton (mimeograph) suggest that nouns may be easier to learn than other parts of speech because they are more imageable (see *5*, below).

2. Distinctiveness of word form

A word that has a distinct form is easier to learn than a word that is similar in form to some other word. Nation (1990) describes the problems he had distinguishing two Indonesian words, *bintang* (star) and *binatang* (animal). Huckin and Bloch (1993) in a qualitative study of contextual inferencing found that "mistaken identity" accounted for most of the cases of unsuccessful guessing. The learners allowed word shape to override contextual factors. For example, one learner misread *optimal* as *optional*. In both of these examples, it is interesting to note that the words causing the problem begin and end with the same letters, being differentiated by letters in medial position. Operating principles

may be involved here also. Where oral input is concerned, mistaken identity, where unfamiliar words are heard as familiar words, is likely to be even more common. According to Ellman and McClelland's (1984) trace model of lexical access, lexical expectations can override an auditory analysis of the input.

3. Length of word form

It seems reasonable to suppose that learners find it easier to decode and remember monosyllabic than polysyllabic words. Meara (1984) reports that Chinese learners of L2 English were found to have "an unexpected difficulty with long words" (p. 234). Ellis (forthcoming) reports that Japanese learners of L2 English found it much easier to recall new short words they had been exposed to in classroom input than new long words. Meara suggests that the reason for the Chinese learners' difficulty might lie in the lack of familiarity with the Roman alphabet (i.e., a reading problem). However, Ellis' learners were exposed to oral input only and yet also performed badly on polysyllabic words like *cannister*. Once again, the explanation probably lies in the nature of the processing strategies learners use. It seems very likely that more effort and time are needed to process and remember polysyllabic words than monosyllabic words.

4. Degree of correlation between form and meaning

There is considerable evidence to suggest that learners initially store words according to their sound rather than their meaning. Henning (1973, cited in Nation, 1990) found that low-proficiency learners tended to connect similar sounding words such as *horse* and *house* whereas high-proficiency learners made connections according to meaning (e.g., *horse* and *cow*). Similarly, Meara's (1982) word association experiments showed that low-level learners were predisposed to make clang associations, whereas high-level learners and native speakers made more paradigmatic associations. If learners are initially strongly influenced by the sound of words, then it would seem likely that they will find it easier to learn words where there is a degree of correlation between sound and meaning (onomatopoeic words like *bark* or *hiss*) and also, perhaps, words where the form clearly signals the grammatical meaning of a word (na*tion*; lov*able*).

5. Imageability

Ellis and Beaton (mimeograph) review research which indicates that imageable words are more easily learned than non-imageable words. An imageable word is one that "arouses a mental image" (p. 5). Concrete words like *dog* and *cup* may be more easily learned than abstract words like *love* and *cruelty*, because they are more imageable.

Although the studies they report (including their own) concern intentional learning, it would seem highly likely that imageability is an important factor in incidental learning too, probably because, as Ellis and Beaton note, it results in a "richer representation." Elley (1989), in the study referred to above, found that vividness was one of six factors that was significantly correlated with his 8-year-old subjects' mean vocabulary gains, although less strongly than some of the other factors he investigated.

Input Factors

Long (1981) distinguishes *input* and *interactional* factors in discourse involving learners. Input factors are formal in nature; they reflect the ways in which the speech addressed to L2 learners is simplified. Interactional factors are functional in nature; they reflect the ways in which discourse involving L2 learners is modified. The distinction is not an absolute one, of course, as interactional modifications contribute to input simplifications. The distinction is motivated primarily by the fact that input simplification can take place without any interactional modification.[6]

1. Frequency

An input factor of obvious importance is word frequency. There is plenty of evidence to suggest that the number of encounters learners have with a word is a major factor in whether they learn it or not. Most of the research, however, has addressed frequency in written texts. Kachru (1962, cited in Nation, 1990) found that words that occurred seven or more times in the coursebooks of Indian learners were known by most learners, but that most of the words occurring only once or twice were not known by most learners. Palmberg (1987) found that the vocabulary remembered by beginner learners of English in a classroom context in Sweden reflected the frequency of the items in the textbook. Beck, McKeown, and Omanson (1987) report a vocabulary training study, part of which investigated the effects of word frequency on vocabulary learning by native speaking children. They found that four encounters with new words were sufficient to produce some results, although the learning that took place was "limited." Elley (1989) found that frequency produced the highest correlation with vocabulary gains made by 7- and 8-year-old children from stories read to them. Eller, Pappas, and Brown (1988), in another study investigating the effects of listening to stories, also found frequency a factor. Brown (1993) found a clear, although not very strong, relationship between general word frequency (determined by reference to the Brown University corpus, Francis & Kucera, 1982) and the words learned by learners enrolled in a

university intensive English language program during one semester. However, she did not find any relationship between the specific vocabulary used in a video-instructional program, "Raiders of the Lost Ark," and vocabulary learning.

A question of considerable importance is whether the relationship between frequency and word learning is a monotonic one (i.e., the more frequent a word is, the more learners learn it and the greater the depth of their learning) or whether there is a threshold effect operating (i.e., a word has to be encountered a specific number of times for it to be learned). The results to date suggest that little is learned from just one or two exposures. Hulstijn (1992), for instance, notes that retention of word meanings in an incidental learning task that afforded just one exposure was "very low indeed." Brown's failure to find any significant effect for specific vocabulary frequency may reflect the fact that the words did not occur with sufficient frequency for any measurable learning to take place. In contrast, her finding that general vocabulary frequency was related to learning can be explained by the fact that the learners in her study may have received sufficient exposures for learning to take place. This points to a threshold effect. It should be noted, however, that Nagy, Herman, and Anderson (1985) argue that a small increment in word knowledge is possible from a single exposure to written text.

In fact, much depends on the depth of knowledge being measured. It is possible that a single exposure is sufficient for some learning to take place, even in the case of oral input. Learners, for example, may be able to recognize it as a "real" word where previously they could not. But a single encounter is unlikely to be sufficient to enable learners to give the meaning of a word, let alone recall its phonological form.

If by "knowing a word" we mean being able to give its meaning, then the available research suggests around six or seven exposures seem to be the minimum necessary, although this is obviously likely to vary depending on various other factors, such as the intrinsic difficulty of the word.[7] Palmberg (1987) refers to research by Carpay (1975) which indicates that new words should be used in at least four different contexts when they are first introduced, followed by subsequent exposure on three other occasions. The assumption underlying the "threshold" view is that each exposure adds something to a learner's knowledge of a word, moving it "a little bit higher on the scale of knowledge" (Nagy & Hermann, 1987, p. 27), but that not until there have been sufficient exposures will learners "know the meaning" of a word.

2. Saliency through focus

A word may be salient (noticeable) because of its intrinsic properties, or it may be made salient by the way it is used in the input. One way that words can be made salient in oral input is through repetition (i.e., frequency), but there are other ways. A word can receive emphatic

stress. A word may be a key word for understanding a particular message. A word can be topicalized through left dislocation (e.g., "The dishwashing liquid—give it to me"). These devices serve to focus the learner's attention on the word and thus make learning more likely.

One way in which words can be made salient is through instructional focus. This may or may not involve making them the subject of intentional learning. For example, Hultsijn (1992), in the study referred to earlier, asked learners to read a passage containing unknown words, clues to the meaning of which were provided in the margins. Hulstijn argues that this still constituted an incidental learning task because the learners were concerned only with understanding the text, not learning the words. Hulstijn found that such focused exposure had some limited impact on learning. Brown (1993) also examined saliency through instructionally-focused exposure by investigating whether glossing words and using them in exercises based on the "Raiders" video affected learning. It did not. However, she found that "salience from gap" affected learning. This referred to the idea of the concept labeled by a word being introduced before the word itself (e.g., the learners might see a shovel being used in the "Raiders" film, but not hear or see the word shovel until later). Brown concluded that "words that are important in a specific context are more likely to be acquired regardless of frequency" (p. 281). Elley (1989) found that words that were helpful to the plot and characterization were more likely to be learned from listening to stories than words that were less important in these respects.

3. Availability of contextual cues

There is general agreement that contextual cues aid vocabulary acquisition. Elley (1989), for example, found that strength of meaning cues strongly correlated with vocabulary gains in the learners he investigated.

Sternberg (1987) provides a useful classification of the different types of contextual cues found in reading texts. As these are also applicable to oral input, they are summarized in *Table 1*. He points out, sensibly enough, that "it is unrealistic to expect a given kind of cue to be equally helpful in figuring out the meanings of all kinds of words" (p. 93). Some cues will work better for some words and other cues for other words (e.g., a temporal cue for *diurnal*; or a spatial cue for *pasture*, as in *low lying pasture*). He also notes that the helpfulness of a cue can be determined by whether it precedes or follows the word. Studies of cue utilization in reading suggest that it is easier to make use of preceding cues than of following cues (e.g., Huckin & Bloch, 1993) although this might not be the case in oral input, as learners might benefit from first hearing the unknown word and then receiving clues to its meaning. Finally, the proximity of the cue to the unknown word is also likely to be important. Huckin and Bloch found that in most cases of successful guessing their subjects used some word or words in "the immediate context."

It might be argued that input which provides learners with rich clues to the meaning of an unknown word is more likely to facilitate acquisition of its meaning than input that provides only poor clues. This is highly likely where intentional learning is concerned. Li (1988) compared the effect of "cue-adequate and cue-inadequate contexts" on the inferencing and retention of unknown words from both written and oral input by 48 advanced Chinese learners of English. In this study, the learners were presented with "60 discrete, semantically disconnected sentences" (p. 405), so it seems highly likely that they approached it as a task calling for intentional learning of the items. The results showed that both inferencing and retention were facilitated by cue-adequate contexts. However, as we noted earlier, it is not at all clear that this will also be the case in incidental vocabulary learning. A rich context may make understanding possible without any learning taking place.

There is clearly a need for much more research on the effects of contextual cues on vocabulary learning from oral input. We know little about the effect of different types of cues, the proximity of the cue to the unknown word, and cue-adequate as opposed to cue-inadequate contexts. It ought to be possible to tease out these effects through careful experimentation. We also need to go beyond the micro-level by examining how macro-structural properties of the context influence acquisition. For example, we might examine how learners make use of their knowledge of "story grammar" to learn new words.

4. Input complexity

The overall complexity of the input in which new words are embedded can also be expected to effect their learnability. There is a considerable literature on the relationship between input complexity and comprehension (see Ellis, 1994). Parker and Chaudron (1987), for example, review 12 experimental studies of the effects of input modifications on comprehension. Their conclusion is that although inguistic modifications (e.g., simpler syntax and vocabulary) helped comprehension, they did not do so consistently. In contrast, what they call *elaborative modifications* (e.g., repetition, paraphrase, use of synonyms, extraposition, and cleft constructions) were considered more likely to promote comprehension. However, Parker and Chaudron's own study failed to show any advantage for elaborative modifications, a result they explain by suggesting that the overall lexical and syntactic difficulty of the reading texts they used was so great as to negate the effects of the modifications. This research is premised on the assumption that if input aids comprehension it also aids learning; but, as we have already noted, this assumption may not be fully justified.

TABLE 1
Types of Contextual Cues (from Sternberg, 1987)

TYPE OF CUE	DESCRIPTION
temporal	Cues regarding the duration or frequency of X (the unknown word) or when X can occur.
spatial	Cues regarding the location of X, or possible location in which X can sometimes be found.
value	Cues regarding the worth or desirability of X, or the kinds of affects X arouses.
stative descriptive	Cues regarding properties of X (such as size, shape, color, odor, feel etc.).
functional descriptive	Cues regarding possible purposes of X, actions X can perform, or potential uses of X.
causal enablement	Cues regarding possible causes of, or enabling conditions of, X.
class membership	Cues regarding one or more classes to which X belongs, or other members of one or more classes of which X is a member.
equivalence	Cues regarding the meaning of X, or contrasts (such as antonymy) to the meaning of X.

Two measures of input complexity that can be expected to have an impact on both the comprehension and acquisition of new words are *lexical density* and *density of unknown words*. Lexical density is determined by establishing the ratio of content words to total word tokens. Faerch, Haastrup, and Phillips (1984, p. 84) comment: "The underlying assumption is that, as it is the content words which primarily convey information, a text is dense if it contains many content words relative to the number of function words."

The density of unknown words is measured by establishing the ratio of new words to total words in a text. Writers of simplified reading texts set considerable store in controlling the number of new words learners will encounter. Nation (1990) refers to a study by Holley (1973) which tried to determine the optimal level of new-word density for written texts. It indicated that "vocabulary learning continues to increase even up to a new vocabulary density of one new word per 15 known words" (p. 343). This is considerably denser than the 1 to 50 density recommended by West (1955) for supplementary readers. Nation considers that Holley's learners may have been able to handle a fairly high density because the overall length of the text they were asked to read was fairly short (750 words). Of course, the optimal lexical density and new-word density is likely to be lower for oral texts than for written texts. There appears to be little research addressing this phenomenon.[8]

Interactional Factors

As we have already noted, what distinguishes oral input from written input is that the latter typically involves interaction. This allows for the input to be modified at just those points where learners experience comprehension problems. When learners hear a word that prevents them from understanding the message, they can request clarification and thereby obtain further clues about its meaning. Interaction, therefore, potentially enriches input in a number of ways that can be hypothesized to facilitate word acquisition.

The literature on meaning negotiation in interactions involving L2 learners is extensive (Long, 1981; Gass & Varonis, 1985; Pica & Doughty, 1985; Ehrlich, Avery, & Yorio, 1989; Gass & Varonis, 1991) but has focused on the nature of the modifications that take place, the structure of the modified discourse, and on comprehension. There has been no attempt to examine the nature of what Gass and Varonis (1985) refer to as the *trigger* (the source of the communication problem). However, a quick look at the protocols given as examples of meaning negotiation in the above articles makes it clear that, in many instances, the trigger is lexical in nature. Often the requests for clarification focus

quite explicitly on word meaning, as in this example from Pica and Doughty (1985):

A: *She's on welfare.*
B: *What do you mean by* welfare*?*

However, while there is ample evidence that this kind of meaning negotiation improves comprehension of the overall message, and while there are also strong theoretical grounds for believing that it promotes acquisition (see Long, 1981) there is, as yet, little empirical evidence that it helps learners learn new words. Ellis, Tanaka, and Yamazaki (forthcoming), however, report a study which suggests that access to interactionally modified input results in more words being learned than access to premodified input. How then does interaction help learners to acquire vocabulary? Two possible ways are (a) by increasing the overall quantity of the input, and (b) through elaboration that results in greater transparency of meaning.

Interaction where meaning negotiation takes place is likely to result in more input. Ellis, Tanaka, and Yamazaki (forthcoming) found that teachers' directions which learners were allowed to negotiate were considerably longer than directions addressed by native speakers to other native speakers and also longer than directions that had been specially prepared to facilitate the learners' comprehension. It seems reasonable to suppose that the more input learners get, the faster they will learn.[9] Where vocabulary is concerned, sheer quantity of input may help by giving learners more time to process new words. Perhaps they tune out the redundant input, thus creating time to focus on the items they find problematic.

Gass and Varonis (1985) define *elaboration* as information not included in the original problem-causing utterance, but subsequently added. One of the main ways in which speakers elaborate on their initial input when the problem is lexical is by adding a definition of some kind, as in this example from Ellis, Tanaka, and Yamazaki (forthcoming). In this study, students were exposed to instructions that required them to place pictures of various kitchen objects in a matrix picture of a kitchen.

T: *We have an apple. And I'd like you to put the apple in the sink.*
S: *What is the sink?*
T: *Sink is a place to wash dishes. It's a hole where you wash dishes.*
S: *One more time please.*
T: *We have an apple. And I'd like you to put the apple in the sink.*
S: *What is sink?*
T: *Sink is a hole and you wash dishes in the hole.*

In this sequence the learners are exposed to a word they do not know—*sink*. A request for clarification elicits a quite explicit definition from the teacher. In fact, the definition meets all the requirements set by Aristotle, as summarized by Edwards (1967, p. 322):

1. A definition should give the essence or nature of the thing defined, rather than its accidental properties.
2. A definition should give the genus and differentia of the thing defined.
3. One should not define by synonyms.
4. A definition should be concise.
5. One should not define by metaphors.
6. One should not define by negative terms or by correlative terms.

It seems likely that learners will obtain a very good idea of the meaning of words so defined. Ellis, Tanaka, and Yamazaki's study was also able to show that three-quarters of the learners exposed to this definition did in fact learn the meaning of *sink*.

Elaboration through definitions is also found in non-interactive discourse (e.g., lectures). However, such elaboration does not always result in the same degree of transparency evident in the above example. Chaudron (1982) describes some of the devices used to elaborate in teacher-talk. He argues that some of them (apposition, conjunction, and parallel structures) can result in ambiguity; learners cannot tell "whether new meanings are being added or whether a particular word is being elaborated" (p. 16). Here is one of Chaudron's examples taken from a high school 10th-grade history lesson. It shows how apposition might confound rather than facilitate word acquisition:

> *The beaver is known as a very industrious and busy, uhm, hard-working animal.*

The learners will find it impossible to know whether *busy* and *hard-working* are synonyms of *industrious*, or new ideas. One of the advantages of negotiated interaction is that ambiguous definitions of this kind are less likely to occur.[10]

Interaction is an important factor in vocabulary acquisition from oral input. So far, the research has focused on the role of negotiation centered on lexical problems arising in communication. The role of interaction in vocabulary acquisition, however, goes way beyond meaning negotiation. We need to broaden the scope of inquiry by examining how interaction helps learners to focus attention on specific items and to clarify their meanings, even when there are no problems. We should also recognize that interaction is not a necessary condition of acquisition—learners can acquire vocabulary from non-interactional input, as studies like Elley (1989), referred to earlier, have shown.

Learner Factors

An explanation of how incidental vocabulary acquisition takes place must take account of the learner as well as the input. Any model of vocabulary acquisition that leaves out learner factors is incomplete. From the existing literature, we can identify a number of learner factors that appear to be important.

1. Background knowledge

A crucial factor is the learner's background knowledge. It is widely acknowledged that there is a close relationship between vocabulary breadth and world knowledge. As Drum and Konopak (1987) put it, "the more knowledge of the world that an individual has, the more words that person will know." Given that world knowledge is an important component of general intelligence, vocabulary can be expected to provide a good indicator of a person's overall intelligence, a hypothesis supported by psychometric IQ tests (Sternberg, 1987). What, then, is the nature of the relationship between what a person knows and word acquisition? The answer is almost certainly that it is an interactive one. Learners with wide background knowledge will find it easier to infer the meanings of new words from context and thus increase their vocabulary. Conversely, learners with a wide vocabulary will find it easier to utilize contexts in the acquisition of new concepts.

Imagine an L2 learner hearing (or reading) the following:

> *The earthquake was devastating. It destroyed half the buildings in the city. Furthermore, the resulting* tsunami *flooded the whole area.*

There are clues within the context as to the meaning of *tsunami* but these are likely to be insufficient without some background knowledge (e.g., that earthquakes disrupt the seabed, causing tidal waves). The learner who already has some knowledge of how an earthquake can cause a tidal wave is in a much better position to understand and perhaps to learn *tsunami*.

So too is the learner with a wide existing L2 vocabulary. This is true on two accounts. First, without a knowledge of the meanings of *earthquake*, *resulting*, and *flooded*, the learner will not have access to the contextual cues that signal the meaning of *tsunami*. Coady (1993) claims that words that are easily recognized serve as "access portals to schemata" (p. 11). Second, if the new word is to be learned, the learner has to slot it into an existing lexical network. Aitcheson (1987) has suggested that a mental lexicon consists of overlapping modules (one for semantic-syntactic information, one for phonological information, and one for creating new words). Each module is characterized as a complex network and within each network "there should be clusters of dense,

multiplex mini-networks" (p. 198). Aitcheson suggests that one kind of mini-network is composed of words related to the same semantic domain. The richer the network representing a domain, the easier it is to make a new entry. Thus a learner with a tightly organized network that already includes such items as *earthquake*, *devastating*, and *flood* is better equipped to learn *tsunami*.

Learners also possess *word schemas*. Nagy, Anderson, and Herman (1987); Nagy and Scott (1990); and Nagy and Gentner (1990) identify various types of word knowledge that can help a learner learn new vocabulary. *Morphological knowledge* consists of knowledge of the prefixes, roots, and suffixes that make up words. *Knowledge of the typical patterns of word meaning* constrains the range of hypotheses that learners have to entertain when they encounter a new word. Learners have knowledge of basic semantic relationships such as synonymy, antonymy, and hyponymy. They also have knowledge of the properties of semantic domains (e.g., they know that English verbs like *slide* and *wobble* say something about how an object moves but not about the shape of the object that is moving). Nagy and his co-workers hypothesize that learners use these different kinds of knowledge to construct and evaluate hypotheses about the meanings of new words. In one study (Nagy & Scott, 1990) for example, they show that 7th-graders' ability to distinguish between plausible and implausible English word meanings was related to the size of their lexicons.

From this discussion of the role of background knowledge we can suggest the following general scale of learning difficulty:

Difficult: The learner does not have an existing concept to which the L2 item can be attached; the lexical domain to which the item belongs is underdeveloped; the learner's word schemas are poorly developed.

The learner has an existing concept to which the L2 item can be attached; the lexical domain to which the item belongs is well developed; the learner's word schemas are well-developed.

Easy:

Again though, we need to acknowledge that the fact that a word is easy to learn with the help of background knowledge does not guarantee that it will be learned.

2. Procedural knowledge

Learners also have metacognitive and metalinguistic knowledge about how new words can be learned. This kind of procedural knowledge is defined by Faerch and Kasper (1985) as "knowing how." It consists of

the strategies and procedures employed by learners to process L2 data for acquisition and use. In the case of vocabulary acquisition we can distinguish two types of procedural knowledge: that associated with contextual inferencing, and that associated with negotiating meaning in interaction.

A number of researchers have examined the procedural knowledge involved in contextual inferencing, but probably the most substantial work is that of Haastrup (1991, 1992). Haastrup's model is based on her analysis of the strategies evidenced by Danish learners of English in think-aloud tasks which required them to work in pairs to guess the meanings of unknown words from written texts. Haastrup describes a continuum of word-processing strategies, reflecting the different types of cues that learners utilize. At one end of the continuum is "pure bottom-level processing," where learners rely exclusively on word-internal cues, while at the other end is "top-level processing with full integration," where learners use a variety of cues taken from different levels, including context and semantics, and integrate them. Haastrup (1991) suggests that there is a "developmental continuum" of processing ability. Learners at the first stage of development can process only zero items (words with no available linguistic cues such as *bouts*). At the second stage they can process *ling* items (items like *curative* which invite interactive processing with integration of just one linguistic element). At the third stage they can handle *ling+* words (words that invite interactive processing with integration of several linguistic cues, e.g., *indiscriminately*). Although Haastrup's claims about a developmental continuum are somewhat speculative, her research provides strong evidence of variation in the individual learner's ability to process meaning from context and also points to the way in which this ability might develop over time.

Whereas Haastrup is concerned with the processing knowledge involved in inferring meanings from written contexts, Robinson (1993) considers the procedural knowledge involved in negotiating meaning, in particular lexical meaning. He notes that negotiation involves the twin processes of *assertion* (one participant trying to fix or explain a word meaning) and *assimilation* (the other participant trying to break down what is said into more manageable units in order to assimilate them). He argues that "procedural vocabulary is essential to this process" (p. 241). High frequency vocabulary—Robinson calls it *indexical vocabulary*—serves as the means by which learners negotiate understanding of new words.

Obviously, learners vary in their procedural knowledge. We can expect learners with well-developed abilities to carry out inferencing involving the integration of varied cues, linguistic and contextual/semantic; and learners who are adept at negotiating the meanings of new words to develop their lexicons—both in terms of breadth and

depth—more rapidly than learners whose skills are less developed in these respects.[11]

3. Immediate phonological memory

In order to learn new words it is necessary to form a stable representation of their phonological form. The extent to which individual learners are able to form phonological representations rapidly and accurately, therefore, is likely to influence their ability to acquire L2 vocabulary. In the case of L1 vocabulary acquisition, Gathercole and Baddeley (1990) review research which suggests there is a close link between children's phonological memory skills and vocabulary development. They also report a study which provides results suggesting that 5-year-old children with poor ability to repeat non-words were slower at learning unfamiliar names for toys than children with good repetition skills. It would seem highly likely that immediate phonological memory is a significant factor in L2 vocabulary acquisition. It is also possible that phonological memory improves as a result of acquiring new words.

4. The learner's L1

There is ample evidence to suggest that the learner's L1 plays a major role in L2 vocabulary acquisition. One important factor is the distance between the L1 and L2. Where the two languages are close (e.g., English and Dutch), learners will have access to a large number words cognate with L2 words which have the same or very similar meanings. In such cases, learners are likely to find the acquisition of both the phonological and semantic representation of a "new" word easy. Nagy et al. (1993) report a study of Spanish-English bilinguals and biliterates, which shows that they were effectively able to transfer Spanish lexical knowledge to reading in English. Even languages that are not close may contain borrowings that facilitate acquisition. Japanese, for example, contains a large number of words borrowed from English. However, because the pronunciation of these words has been adapted to the L1 phonological system (a process known as *foreignizing*) learners may have difficulty in recognizing the L2 words when they hear them.

Also, care needs to be taken not to over-emphasize the value of cognates to learners. The fact that an L1 word has an L2 equivalent where its basic, denotive meaning is concerned does not permit the learner to conclude that the two words are identical with respect to other aspects of use and meaning—the grammatical patterns in which they can be used, their connotation, collocation, and association. Holmes and Ramos (1993) note that cognates can often prove "false friends." They point to the need for learners to check their initial guesses based on the recognition of cognates against other contextual cues. Nagy et al. (1993) suggest that the ability to distinguish true from false or partial cognates

TABLE 2
A Summary of Factors Potentially Influencing the Acquisition of L2 Words

FACTORS	DESCRIPTION
intrinsic word properties	
part of speech	Learners learn nouns more easily than other parts of speech.
distinctiveness of word form	A word with a distinct word form is learned more easily than a word which is similar in form to a word the learner already knows.
length of word form	Monosyllabic words are easier to learn than polysyllabic words.
correlation between form meaning.	Learners find it easy to learn words where there is a degree of correlation between sound and and meaning.
imageability	A word that arouses a mental image (e.g., a concrete noun) is easier to learn than a word for which there is no obvious mental image (e.g., an abstract noun).
input factors	
frequency	Words encountered frequently are learned more readily than words encountered rarely; also, frequent encounters promote 'depth' of learning. The relationship between frequency and vocabulary learning may not be monotonic, however.
saliency through 'focus'	Words that are made salient through some kind of focus (e.g., instruction) are more likely to be remembered.
availability of contextual cues	Words for which there are cues in the immediate context are likely to be more easily learned than words for which there are no contextual cues, or only distant cues. But currently little is known about what type of cue works best.
input complexity	New words that occur in simplified input (i.e., input with low lexical density and low new-word density) are likely to be more readily learned than words that occur in more complex input.

FACTORS	DESCRIPTION
interactional factors	
resulting in more input	Interaction can increase the total input available to learners,which may give them more time to process items they find problematic.
resulting in elaborated output	Interaction enables learners to request clarification and thereby obtain definitions of unknown words. However, not all elaborated input is beneficial for acquisition.
learner factors	
background knowledge	Background knowledge helps learners to inference the meaning of new words. Possession of a richly specified lexical domain makes it easier to 'place' a new word in the domain.
procedural knowledge	Learners vary in their ability to integrate bottom-up and top-down processing and, therefore, in their ability to inference the meanings of new words. Learners also vary in their knowledge of the procedural vocabulary needed to negotiate understanding of new words.
immediate phonological memory	Learners vary in their ability to form accurate phonological representations of new words they hear.
the learner's L1	L2 words cognate with L1 words are easy to learn, although they may sometimes prove 'false friends.' L2 words with a similar phonological shape to L2 words may also be easier to learn.

may require "a high level of cognitive flexibility and metacognitive awareness" (p. 254).

The degree of similarity in the phonological features of the L1 and L2 are also likely to affect the ease of learning L2 vocabulary. Ellis and Beaton (mimeograph) review research which indicates that "the less overlap between the feature set of the native and the foreign word, the

harder it will be for the FL learner to learn that word" (p. 3). Thus, a Chinese learner of English finds words like *pen* and *see* easy to learn but words like *rice* and *eighth* difficult.

Conclusion

The purpose of this paper has been to examine the factors that are likely to influence learners' success in learning L2 vocabulary incidentally from oral input. Unfortunately, this important aspect of second language acquisition has been largely neglected by researchers. While there is a substantial body of research relating to vocabulary acquisition from written context, hardly any attention has been paid to oral context. As a result, this paper is programmatic and, in places, speculative. It nevertheless seems important to ask how vocabulary acquisition from oral input takes place—if only because learners attach considerable importance to vocabulary and because many L2 learners are dependent on oral input, particularly in the initial stages of acquisition.

To summarize the main conclusions:

- Most L2 vocabulary is learned incidentally, much of it from oral input.

- L2 vocabulary acquisition needs to be characterized as having a quantitative and a qualitative dimension. Learning a word involves a gradual process of adding depth to an initial, shallow representation.

- A variety of factors are likely to affect the learnability of a word. These have been grouped under four general headings: the intrinsic properties of the L2 word, input factors, interactional factors, and the learner's background knowledge (including knowledge of the L1). The different factors relating to these four categories are summarized for convenience in *Table 2*

Finally, a word of caution. It is necessary to reiterate what has been a common theme throughout this paper—that comprehending and learning an L2 word are not the same thing. Much of the above discussion has centered on what makes comprehension of new L2 items possible, on the assumption that comprehension is a necessary condition for learning to take place. It is not a sufficient condition, however. In fact, comprehension may obviate the need for acquisition; learners may understand a text without the need to notice a new item. This raises the question as to what else is needed for incidental vocabulary acquisition to take place. One answer is interest. Elley (1989) found that the children in his study picked up almost twice as many words from listening to one story than to the other and suggests that the principal reason for this was "lack of involvement" in the second story. Stoller and Grabe (1993, p. 38) argue that "because L2 students often find themselves at a 'loss for words,' they are usually quite motivated to

improve their vocabulary." However, not all learners are so motivated and, more particularly, as Elley's study indicates, they are not necessarily motivated to learn every new word they come across. Ultimately, an account of incidental L2 vocabulary acquisition will have to address personal and affective factors as well as psycholinguistic issues.

Acknowledgment

I would like to thank Stephen Krashen, Leo Van Lier, and Lidia Woytak for their helpful comments.

Notes

1. Van Lier, in his comments on a draft version of this paper, rightly points out that the distinction between incidental and intentional learning is more problematic than my discussion of it acknowledges. He notes, for example, that intentionality may be considered as something external to the learner (e.g., imposed by an instructional focus) or internal (i.e., the learner's own preferred orientation). He also notes that language users alternate between focal or peripheral attention in response to processing needs.

2. Krashen (1989), however, provides evidence to suggest that incidental learning may be more effective than intentional learning. He surveys a number of studies that have investigated the impact of silent reading on vocabulary growth and the relative effects of self-selected reading and regular reading programs. He concludes that "free readers do at least as well, and often better, than students in the regular programs on vocabulary tests, suggesting that free reading is at least as effective as traditional instruction" (p. 448). The contradiction between Krashen and Pressley et al. may be more apparent than real, however. Krashen does not consider short-term experimental studies, which do tend to find in favor of intentional learning. In long-term studies, where incidental learning is given a chance to work, the difference is less marked and, as Krashen demonstrates, often favors incidental learning.

3. While it is true that there is no widely accepted model that explains how learners come to understand the meaning of words they do not know from context, a number of preliminary attempts to construct a model have been made (e.g., Huckin & Bloch, 1993).

4. Krashen, in comments on a draft version of this article, argues that the problem Coady identifies may not be a problem at all. He argues

that a clear context can lead to acquisition even when there is no "need" for the learners to attend to the item. He gives this example:

> If an acquirer of English has not yet acquired the word *door*, but knows the word *close*, and someone asks him to close the door in a very clear context, the acquirer's LAD (language acquisition device) will assign the correct meaning to *door*, and at least some acquisition will take place.

However, Krashen goes on to admit that when there is "extremely rich contextual support" the learner may not notice a new item in the input. Krashen's argument suggests that we need to examine the kind of contextual support that promotes (and inhibits) acquisition. The later section on availability of contextual clues addresses this.

5. Whereas classroom learners are likely to be exposed to large quantities of written input, untutored learners such as those investigated in the European Science Foundation project on adult second language acquisition will be dependent on oral input for developing their lexicons.

6. It is customary to distinguish *premodified input* (input that has been simplified to be appropriate to a particular level of learner) and *interactionally modified input* (input that is modified dynamically through the course of an interaction)—see Pica, Young, and Doughty (1987). Graded readers contain premodified input; face-to-face interaction offers opportunities for interactionally modified input.

7. Nation (1990) refers to a study by Saragi et al. (1978) which shows that in an incidental learning task, 16 or more repetitions in a written text were necessary.

8. Rost's (1990) book, *Listening in Language Learning*, for example, has nothing on lexical density or new-word density as factors influencing input comprehension.

9. Wells' (1985) study of children learning their first language suggests that quantity of input is a major factor in accounting for differences in rate of acquisition. Those children who learned the fastest were the ones that had the most language addressed to them. However, there were also qualitative input differences that were important.

10. It is interesting to note that Drum and Konopak (1987), in a review of research that has investigated the usefulness of different types of cues to word meaning in written text, report that definitions and synonyms have been found to be the most useful.

11. It may also be possible to give training to learners who lack procedural knowledge to help them overcome their limitations. A number of researchers consider the role of "strategy training" in this respect (see, for example, Sternberg, 1987; Graves, 1987; O'Malley, 1987).

References

Aitcheson, J. (1987). *Words in the mind: An introduction to the mental lexicon*. Oxford: Basil Blackwell.

Arnaud, P. & Bejoint, H. (Eds.). (1993). *Vocabulary and applied linguistics*. Basingstoke: MacMillan.

Barnard, H. (1961). A test of P.U.C. students; vocabulary in Chotanagpur. *Bulletin of the Central Institute of English, 1*, 90-100 (cited in Nation, 1990).

Beck, I., McKeown, M., & Omanson, R. (1987). The effects of diverse vocabulary instructional techniques. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition*.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*, 69-84.

Bialystok, E. & Sharwood Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second language acquisition. *Applied Linguistics, 6*, 101-117.

Brown, C. (1993). Factors affecting the acquisition of vocabulary: Frequency and saliency of words. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning*.

Carpay, J. (1975). *Onderwijsleerpsychologie en leergangontwikkeling in het moderne vreedetalen-onderwijs*. Gröningen: Wolters-Noordhoff (cited in Palmberg, 1987).

Chaudron, C. (1982). Vocabulary elaboration in teachers' speech to L2 learners. Studies in *Second Language Acquisition, 4*, 170-180.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Crow, J. & Quigley, J. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly, 19*, 497-513.

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.), *Issues in language testing research*. Rowley, MA: Newbury House.

Curtis, M. (1987). Vocabulary testing and vocabulary instruction. In M. McKeown & M. Curtis (Eds.).

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 895-901.

Davies, A., Criper, C., & Howard, A. (Eds.). (1984). *Interlanguage*. Edinburgh: Edinburgh University Press.

Drum, P. (1983). Vocabulary knowledge. In J. Niles & L. Harris (Eds.), *Searches for meaning in reading: Language processes and instruction.* Rochester, NY: National Reading Conference.

Drum, P. & Konopak, B. (1987). Learning word meanings from written context. In M. McKeown & M. Curtis (Eds.).

Ehrlich, S., Avery, P., & Yorio, C. (1989). Discourse structure and the negotiation of comprehensible input. *Studies in Second Language Acquisition, 11,* 397-414.

Eller, R., Pappas, C., & Brown, E. (1988). The lexical development of kindergarteners: Learning from written context. *Journal of Reading Behavior, 20,* 5-24.

Elley, W. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24,* 174-187.

Ellis, N. & Beaton, A. (mimeograph). Factors affecting the learning of foreign language vocabulary I: Imagery keyword mediators and phonological short-term memory. U.C.N.W. Cognitive NeuroComputation Unit Technical Report.

Ellis, N. & Beaton, A. (mimeograph). Factors affecting the learning of foreign language vocabulary II: Psycholinguistic determinants. U.C.N.W. Cognitive NeuroComputation Unit Technical Report.

Ellis, R. (1984). *Classroom second language development.* Oxford: Pergamon.

Ellis, R. (1990). *Instructed second language acquisition.* Oxford: Blackwell.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (forthcoming). Uptake as a measure of classroom language learning. *English Language Teaching Journal.*

Ellis, R., Tanaka, Y., & Yamazaki, A. (forthcoming). Classroom interaction, comprehension and L2 vocabulary acquisition. *Language Learning.*

Ellman, J. & McClelland, J. (1982). Exploiting the lawful variability of the speech wave. In J. Perkell & D. Klatt (Eds.), *Invariance and variability in the speech wave.* Hillsdale, NJ: Lawrence Erlbaum.

Faerch, C., Haastrup, K., & Phillips, R. (1984). *Learner language and language learning.* Copenhagen: Gyldendals Sprogbibliotek.

Faerch, C. & Kasper, G. (1985). Procedural knowledge as a component of foreign language learners' communicative competence. In H. Bolte & W. Herrlitz (Eds.), *Kommunikation im Sprachunterricht.* Utrecht: Instituut 'Franzen,' University of Utrecht.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. (1993). Effects of

listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly, 28*, 70-79.

Felix, S. (1978). Some differences between first and second language acquisition. In N. Waterson & C. Snow (Eds.), *The development of communication*. New York: John Wiley and Sons.

Francis, W. & Kucera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Gass, S. & Varonis, M. (1985a). Task variation and non-native/non-native negotiation of meaning. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*.

Gass, S. & Madden, C. (Eds). *Input in second language acquisition*. Rowley, MA: Newbury House.

Gathercole, S. & Baddeley, A. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439-454.

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11*, 341-363.

Haastrup, K. (1991). Lexical inferencing procedures or talking about words. Tübingen: Narr.

Haastrup, K. (1992). Approaching learning theory from a different angle: A comprehension perspective on the integrated theory of instructed learning. *Die Neueren Sprachen, 91*, 148-165.

Henning, G. (1973). Remembering foreign language vocabulary: Acoustic and semantic parameters. *Language Learning, 23*, 185-196.

Holley, F. (1973). A study of vocabulary learning in context: The effect of new-word density in German reading materials. *Foreign Language Annals, 6*, 339-347.

Holmes, J. & Ramos, R. (1993). False friends and reckless guessers: Observing cognate recognition strategies. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Huckin, T. & Bloch, J. (1993). Strategies for inferring word meaning in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Huckin, T. Haynes, M., & Coady, J. (1993). *Second language reading and vocabulary learning*. Norwood, NJ: Ablex.

Hulstijn, J. (1992). Retention of inferred and given meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics*.

Jamieson, P. (1976). The acquisition of English as a second language by young Tokelau children living in New Zealand (cited in Nation 1990). Unpublished Ph.D. thesis, Victoria University of Wellington.

Kachru, J. (1962). Report on an investigation into the teaching of

vocabulary in the first year of English. *Bulletin of the Central Institute of English, 2*, 67-72.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*, 440-464.

Li, X. (1989). Effects of contextual cues on inferring and remembering meanings of new words. *Applied Linguistics, 10*, 402-413.

Long, M. (1981). Input, interaction, and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences*, 379.

McKeown, M. & Curtis, M. (Eds.). (1987). *The Nature of Vocabulary Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts, 13*, 221-246.

Meara, P. (1982). Word associations in a foreign language: A report on the Birkbeck vocabulary project. *Nottingham Linguistic Circular, 11*, 29-37.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies et al. (Eds.), *Interlanguage*.

Meara, P. (1992). Network structures and vocabulary acquisition. In P. Arnaud & H. Bejoint (Eds.).

Nagy, W., Anderson, R., & Herman, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, 237-270.

Nagy, W., Garcia, G., Durgunoglu, A. & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Reading Behavior, 25*, 241-259.

Nagy, W. & Gentner, D. (1990). Semantic constraints on lexical categories. *Language and Cognitive Processes, 5*, 169-201.

Nagy, W. and Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis (Eds.).

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nagy, W. & Scott, J. (1990). Word schemas: Expectations about the form and meaning of new words. *Cognition and Instruction, 7*, 105-127.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House/Harper Row.

O'Malley, J. (1987). The effects of training in the use of learning

strategies on acquiring English as a second language. In A. Wenden & J. Rubin (Eds.), *Learner strategies for language learning.* Englewood Cliffs, NJ: Prentice-Hall.

Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *Studies in Second Language Acquisition, 9,* 201-219.

Parker, K. & Chaudron, C. (1987). The effects of linguistic simplifications and elaborative modifications on L2 comprehension. *University of Hawaii Working Papers in ESL, 6,* 107-133.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21,* 737-758.

Pica, T. & Doughty, C. (1985a). Input and interaction in the communicative language classroom: A comparison of teacher-fronted and group activities. In S. Gass & C. Madden (Eds.).

Pressley, M., Levin, J. & McDaniel, M. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. McKeown & M. Curtis (Eds.).

Quinn, G. (1968). The English vocabulary of some Indonesian university entrants (cited in Nation, 1990). IKIP Kristen Satya Watjana, Salatiga.

Richards, J. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10,* 77-89.

Robinson, P. (1993). Procedural and declarative knowledge in vocabulary learning: Communication and the learner's lexicon. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Rodgers, T. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *IRAL, 7,* 327-343.

Rost, M. (1990). *Listening in language learning.* London: Longman.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129-158.

Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development.* New York: Appleton-Century-Crofts.

Stein, M. (1993). The healthy inadequacy of contextual definition. In T. Huckin, M. Haynes, & J. Coady (Eds.).

Sternberg, R. (1987). Most vocabulary is learned from context. In M. McKeown and M. Curtis (Eds.).

Stoller, F. and Grabe, W. (1993). Implications for L2 vocabulary acquisition and instruction from L1 vocabulary research. In T. Huckin, M. Haynes & J. Coady (Eds.).

Twadell, F. (1973). Vocabulary expansion in the TESOL classroom. *TESOL Quarterly, 7,* 61-78.

Wells, G. (1985). *Language development in the pre-school years.* Cambridge: Cambridge University Press.

West, M. (1955). *Learning to read in a foreign language* (2nd ed.). London: Longman.

Wode, H., Rohde, A., Gassen, F., Weiss, B., Jekat, M., & Jung, P. (1992). L1, L2, L3: Continuity vs. discontinuity in lexical acquisition. In P. Arnaud & H. Bejoint (Eds.).

Yoshida, M. (1978). The acquisition of English vocabulary by a Japanese-speaking child. In E. Hatch (Ed.), *Second language acquiition.* Rowley, MA: Newbury House.

Author

ROD ELLIS, Professor of TESOL, College of Education, Ritter Hall, Philadelphia, PA 19122. Specializations: second language acquisition research and its utilization in teacher education and language teaching methodology.

Applied Language Learning
1994, Vol. 5, No. 1, pp. 33-71

Intake Factors and Intake Processes in Adult Language Learning

B. Kumaravadivelu
San Jose State University

Synthesizing theoretical and empirical insights from second language acquisition, cognitive psychology, information processing, schema theory, and parallel distributed processing, this paper proposes an interactive framework of intake processes. It identifies intake factors (Individual, Negotiation, Tactical, Affective, Knowledge, and Environmental) *and intake processes (linguistic processes of grammaticalization and language transfer, and cognitive processes of inferencing, structuring, and restructuring) and argues that these factors and processes interweave and interact in a synergic relationship, each shaping and being shaped by the other. According to this interactive framework, input can be converted into intake only if the intake factors and processes are optimally favorable and if a consistent absence of one or a combination of these constructs may result in partial learning. Pointing out that current research yields only a limited and limiting view of L2 development because of its narrow focus on individual intake factors and intake processes in isolation, this paper emphasizes the need to reframe our research agenda in order to address the synergic relationships between and within intake factors and processes.*

Introduction

It is widely recognized that there is a mismatch, both qualitative and quantitative, between the language output produced by second/foreign language (L2) learners and the language input to which they are exposed. In a seminal paper, Corder (1967) highlighted this mismatch and made an important distinction between *input* and what he called *intake*. Since then, several attempts have been made to explore the relationship between input, intake, and L2 development[1] (see,

© 1994, B. Kumaravadivelu

among others, Faerch & Kasper, 1980; Krashen, 1981; Chaudron, 1985; Gass, 1988; Spolsky, 1989; van Lier, 1991). In spite of a quarter century of exploration, we have hardly reached a consensus on the fundamental characteristics of intake, let alone an understanding of the psycholinguistic processes governing it—a state of affairs that attests to the complexity of the construct we are wrestling with. Continuing the exploration, I take a critical look at the concepts of input, intake, intake factors, intake processes, and output, as they relate to adult L2 development in formal contexts, and then attempt to design what I call an interactive framework of intake processes. I do so by building on work already done, and by synthesizing theoretical and empirical insights derived from interrelated areas such as second language acquisition, cognitive psychology, schema theory, information processing, and parallel distributed processing.

Input

Input is operationally defined as oral/written data of the target language (TL) to which L2 learners are exposed through various sources, and which is recognized by them as language input. This definition posits two conditions: availability and recognizability.

The first condition is rather obvious: Input either has to be made available to learners or they have to seek it themselves. One can identify three types of input attributable to three different, but not necessarily mutually exclusive, sources that learners are likely to get/seek input from:

> (a) *interlanguage input*: the developing language of the learners themselves and their peers, with all its linguistically well-formed as well as deviant utterances;
>
> (b) *simplified input*: the syntactically, semantically, and pragmatically simplified language that teachers and other competent speakers use when they talk to L2 learners in and outside the classroom; and
>
> (c) *non-simplified input*: the language of competent speakers without any characteristic features of simplification, that is, the language generally used in the media (TV, radio, and newspapers) and also the language used by competent speakers to speak and write to one another.

Clearly, each of these three sources of input can manifest itself in various modes: spoken/written, monologic/dialogic, formal/informal, and so on.

The second condition is less obvious than the first, but equally important: Input has to be consciously or unconsciously recognized by

learners not only as language input but as something they can cope with. The language data available in the learners' environment has the potential to become usable input when the learners pay attention to it, thereby noticing the mismatch between the speech of competent speakers and their own organization of the TL (Gass, 1988; Schmidt, 1990, 1993; VanPatten, 1990; van Lier, 1991). What actually makes the learners notice and recognize a subset of language exposed to them as potential input is as yet undetermined. Schmidt (1990, 1993) suggests factors such as frequency of occurrence, perceptual salience, linguistic complexity, skill level, and task demands. One might also add factors like learner needs, wants, situations, interests, and motivation.

Intake

Unlike input, the concept of intake has not been easy to pin down. The current L2 literature is replete with varied definitions and myriad explanations for the term intake. Amidst the conceptual multiplicity, one can discern two strands of thought: one that treats intake primarily as product, and the other that treats it primarily as process.[2] Corder, who is credited with formulating the notion of intake, defines it as "what goes in and not what is *available* to go in" (1967, p. 165, his emphasis). Kimball and Palmer (1978, pp. 17-18) define intake as "input which requires students to listen for and interpret implicit meanings in ways similar to the ways they do so in informal communication." This has been echoed by Krashen, for whom "intake is simply where language acquisition comes from, that subset of linguistic input that helps the acquirer acquire language" (1981, pp. 101-102). A common thread running through these definitions is that all of them treat intake as a product, a subset of linguistic input.

There are others who prefer a process-oriented approach to intake. Faerch and Kasper (1980, p. 64), for instance, define intake as "the subset of the input which is assimilated by the IL (interlanguage) system and which the IL system accommodates to." Hatch (1983, p. 81) is in agreement when she defines intake as a subset of input which "the learner actually successfully and completely processed." Likewise, Chaudron (1985, p. 1) refers to intake as "the mediating process between the target language available to the learners as input and the learner's internalized set of L2 rules and strategies for second language development." Liceras (1985, p. 358) also opts for a process-oriented definition when she talks of cognitive capacities that intervene at the level of intake. Gass (1988, p. 206), too, sees intake "as a process of mental activity which mediates between input and grammars."

Notice that the product view identifies intake as a subset of input *before* the input is processed by learners. In other words, intake *is* input, even though it is only a part of it. The process view, however, identifies

intake as what comes *after* psycholinguistic processing. That is, intake is already part of the learner's IL system. According to the product view, intake then is *unprocessed* language input; according to the process view, it is *processed* language input. The two views can be diagrammatically represented as follows:

FIGURE 1
Input, Intake, Output: The Product View

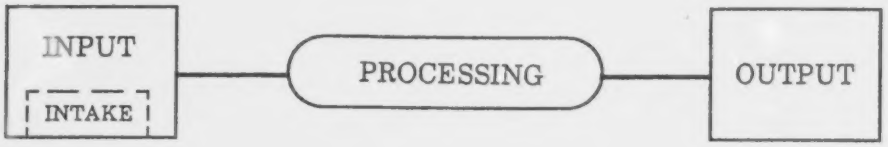
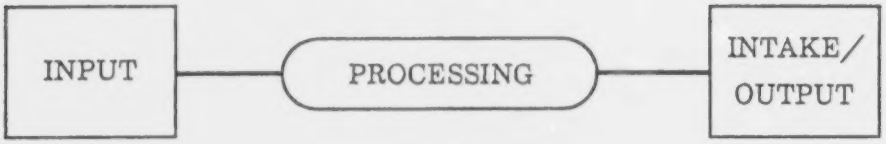


FIGURE 2
Input, Intake, Output: The Process View

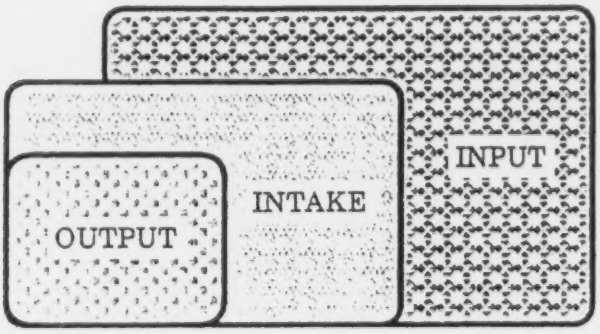


The product view of intake appears to be severely flawed. It implies that there is no need to differentiate input from intake because intake, after all, is no more than a part of input and is independent of language learning processes. The distinction between input and intake, crucial to the nature of L2 development, becomes insignificant if not irrelevant. Furthermore, without such a distinction, we will not be able to account for the fact that "input is not perceived and processed by different learners in an identical manner" (Stern, 1984, p. 393).

The process view of intake is not flawless either. First, this view suggests a simple part-whole relationship between input and intake, and between intake and output. It ignores the fact that "there are parts of a learner's grammar which go beyond the actual input, perhaps because a learner imposes regularities on the data or uses native language markedness values" (Gass, 1988, p. 199). Second, intake is not directly observable, quantifiable, or analyzable; it is a complex cluster of mental representations. What is available for empirical verification is the product of these mental representations. We have a different name for such a product; we call it output.

There is thus a need to redefine the concept of intake. It may be useful to treat intake as an abstract entity of learner language that has been fully or partially processed by learners, and fully or partially assimilated into their developing system. Such an entity is the result of as yet undetermined interaction between input and intake factors mediated by intake processes (see below). This definition suggests that intake is treated as a subset of input only to the extent that it originates from a larger body of language data called input. Features of learners' output can be traced not only to the input they are exposed to but to the dynamics of intake processes as well. Such a view accounts for the fact that the learner's developing system provides instances of grammatically deviant utterances which are not part of input. The relationship between input, intake, and output can be diagrammatically represented as:

FIGURE 3
Input, Intake, Output: A Quantitative View



The figure states that, quantitatively speaking, output is a subset of what has been internalized, which in turn is a subset of input. Further, a small portion of the learner output can go beyond the boundaries of language input. What part of input gets converted into intake appears to be determined by certain factors and processes which I call *intake factors* and *intake processes*.

Intake Factors

Intake factors refers to learner-internal and learner-external factors that are brought to bear on the process of converting a subset of input into intake. Just as scholars differ on the concept of intake, they differ widely on their choice of intake factors as well. Corder (1967, p. 165) suggests that "it is the learner who controls the input or more properly his intake." To learner control he adds "the characteristics of his language acquisition mechanism" as another factor. Corder explains further that "what elements are, in fact, processed from the data that is available is determined by what the current state of the learner's interlanguage grammar permits him to take in at that moment" (Corder, 1978, pp. 81-82). Hatch (1983, p. 80) believes that "if it (input) is held in memory long enough to be processed (or if processing breaks down and the learner asks for a new clarification), it has been taken in." Seliger (1984, p. 45) echoes the same idea: "Long-term memory and its effect on the selection of tactics is what determines when input will become intake."

Krashen (1981 and elsewhere) asserts that comprehensible input and low affective filter are the only two factors which determine intake. He is convinced that "every other factor hypothesized to relate to SLA reduces to input plus low filter" (1983, p. 141). Larsen-Freeman (1983, p. 14), too, suggests that "the key to input's becoming intake is its comprehensibility." Sharwood Smith (1985, p. 402) takes exception to these views and states that it is "particularly unreasonable to give L2 input the unique role in explanation of intake." Instead, he emphasizes the role played by crosslinguistic features in intake processing. According to Swain (1985, p. 236), comprehensible output is crucial for converting input into intake. While these scholars highlight the importance of one or two intake factors which are understandably the focus of their immediate research, Spolsky (1989)—in a comprehensive review of the L2 literature—isolates, defines, and explains no less than 74 factors (he calls them conditions) of varying importance that separately or in combination contribute to L2 development.

The diversity of definitions and interpretations found in the L2 literature is evidently a result of the varied perspectives with which researchers have approached the concept of intake and intake factors. While this multiplicity of perspectives has undoubtedly broadened our understanding of intake, the sheer range of intake factors hypothesized to influence L2 development might hinder meaningful investigation. It seems to me that we need an integrated view of the major intake factors that facilitate L2 development in order to help us make informed judgments about L2 development, and consequently about L2 pedagogy.

The task of isolating major intake factors from a plethora of factors suggested in the literature rests largely on individual perception rather than on indisputable evidence. The latter is in any case hard to come by, in spite of a quarter century of L2 research (see Larsen-Freeman & Long, 1991, and Cook, 1993 for recent reviews). My critical reading of factors that facilitate L2 development has yielded a cluster of six major factors, and two variables within each. Notice that I call these intake factors *facilitating*, not *causal*, factors. I do so because, to my knowledge, no direct causal relationship between any of the intake factors and adult L2 development has been established beyond doubt, nor, as Lamendella (personal communication) points out, would a "causalist view" be worth considering in any case, given our limited understanding of L2 development. It is, however, fairly reasonable to assume that each of these factors plays a facilitating role of varying importance. The major intake factors are:

Individual factors: Age and Anxiety

Negotiation factors: Interaction and Interpretation

Tactical factors: Learning Strategies and Communication Strategies

Affective factors: Attitudes and Motivation

Knowledge factors: Language Knowledge and Metalanguage Knowledge

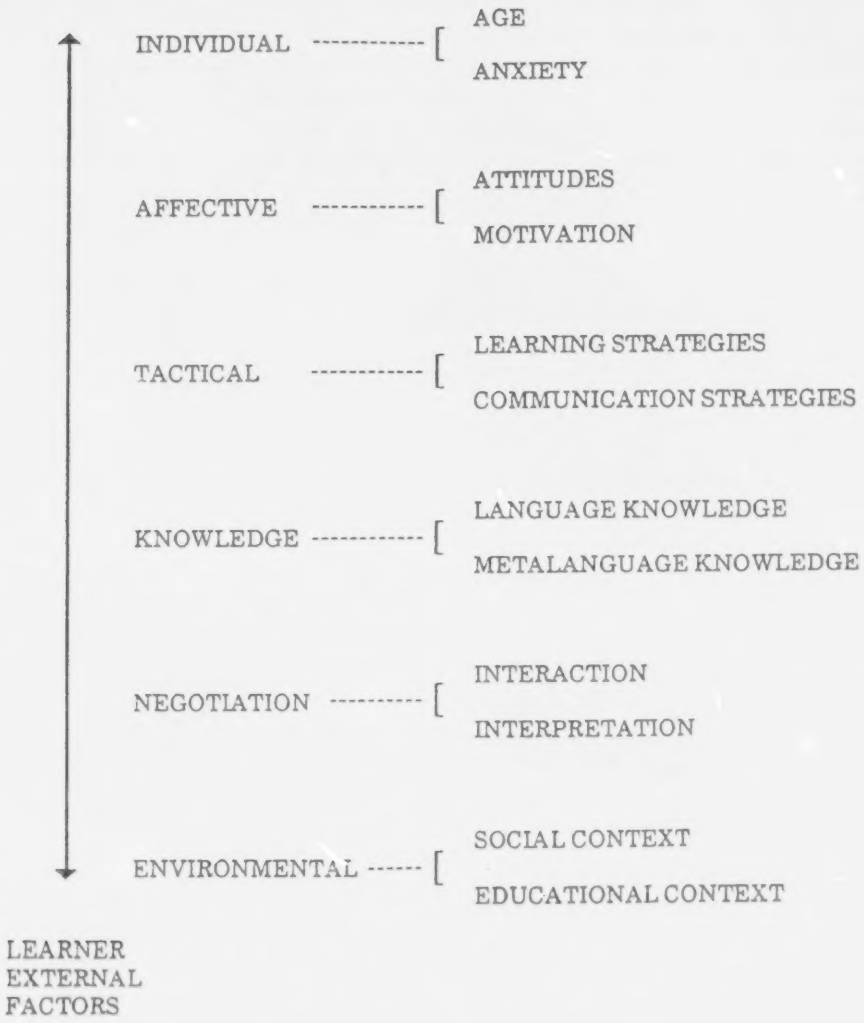
Environmental factors: Social Context and Educational Context

These factors can be classified into two broad categories: learner-internal and learner-external factors. By this categorization, I do not suggest a dichotomous relationship between the two categories; rather, I look at them as the two ends of a continuum as represented in *Figure 4*.

In the rest of this section, I briefly sketch the facilitating role played by each of these intake factors in developing the learner's L2 knowledge/ability.[3] I do so by drawing upon currently available theoretical as well as empirical insights. In the next section (**Intake**

Processes) I shall try to relate the role played by intake factors in activating intake processes.

FIGURE 4
Intake Factors Continuum



Individual Factors

Several individual factors have been studied to assess their role in L2 development. They include age, anxiety, empathy, extroversion, introversion, memory, and risk-taking. Of these variables, age and anxiety appear to play a relatively greater role than the others.

Age

According to Lenneberg's (1967) critical period hypothesis, languages are best learned before puberty, after which everyone faces certain constraints in language development, primarily due to lateralization. While the L2 research based on this hypothesis has yielded mixed results, there seems to be a consensus that a mismatch does exist between the potential for native-like lexical and syntactic knowledge/ability and the potential for native-like phonological knowledge/ability if learning starts after puberty. Native-like accent is almost impossible unless first exposure takes place very early, probably as early as age 6, the reason presumably being that L2 phonological production is the only aspect of language performance that has a neuromuscular basis (Scovel, 1988).

With regard to the development of syntactic and pragmatic knowledge/ability, there are those who suggest that "younger is better" (Krashen, 1981; Johnson & Newport, 1989). Their explanation is mostly based on cognitive capacity—namely, that child and adult L2 development might actually involve different processes, the former utilizing innate properties of language acquisition as in L1 acquisition, the latter employing general problem-solving abilities, and thus accounting for the differential effect of age. But, there are others who suggest that "older is better" because older learners have cognitive and literacy skills which tend to enhance their L2 development (Snow, 1983; Ellis, 1985; McLaughlin, 1987). They suggest that there are contexts in which teenagers and adults not only reach native-like proficiency, but they also progress more rapidly and perform with greater accuracy in the early stages of learning than do their younger counterparts.

A balanced approach suggests a *sensitive* rather than a *critical* period for L2 development (Lamendella, 1977; Singleton, 1989). Such a suggestion acknowledges that certain language skills are acquired more easily at particular times in development than at other times, and that some language skills can be learned even after the critical period, although less easily. It seems reasonable to deduce from research that age will have variable influence on L2 development, depending on which intake factors are brought to bear on the learning experience of an individual learner, when, and in what combination.

Anxiety

Anxiety refers to an emotional state of apprehension, tension, nervousness, and worry, mediated by the arousal of the automatic nervous system. In the context of L2 learning, anxiety is characterized by feelings of self-consciousness, fear of negative evaluation from peers and teachers, and fear of failure to live up to one's own personal standards and goals (Bailey, 1983; Horwitz, Horwitz, & Cope, 1986). Adult L2 learners typically develop a sense of incompetence about internalizing the properties of their L2, and about the inability to present themselves in a way consistent with their self-image and self-esteem.

While psychologists postulate a positive, facilitating anxiety and a negative, debilitating anxiety, each working in tandem (Alport & Haber, 1960), L2 researchers have by and large focused on the effect of the latter. In a series of experiments, Gardner and his colleagues (Gardner, 1985; MacIntyre & Gardner, 1989, 1991; Gardner, Day, & MacIntyre, 1992) found that anxiety has a significant deleterious effect on L2 development. Language anxiety has also been found to correlate negatively with global measures of achievement such as objective tests and course grades as well as measures involving specific processes such as vocabulary recall. Similarly, studies conducted by Horwitz, Horwitz, and Cope (1986), and Madsen, Brown, and Jones (1991) show that a significant level of anxiety is experienced by a majority of their subjects in response to at least some aspects of L2 development. Gardner and his colleagues explain the effects of language anxiety by surmising that it consumes attention and cognitive resources that could otherwise be allocated to developing L2 knowledge/ability. Thus, anxiety may occur at any of the three levels of language development: input, intake processing, or output (Tobias, 1986). At input, it may cause attention deficits; intake processing may be affected because time is divided between the processing of emotion-related and task-related cognition; and, it may also interfere with the retrieval of previously learned information, thereby affecting output. These insights have been supported by diary (Bailey, 1983) as well as experimental (MacIntyre & Gardner, 1991) studies. While a clear picture of how anxiety actually affects L2 development is yet to emerge, it appears that anxiety may have different effects at different stages of L2 development, depending on its interplay with other intake factors and intake processes.

Negotiation Factors

The term *negotiation* has been widely used in ethnomethodology and conversational analysis to refer to the ways in which interlocutors communicate meaning, and structure their social relationships through interaction. Negotiation is important for L2 development because it

implies the use and constant refinement of linguistic perceptions and pragmatic concepts.

There are at least three aspects of negotiation: *introspection*, *interaction*, and *interpretation*. *Introspection* focuses on the particularity of the individual learner. It is intrapersonal, involving, in the Vygotskyan sense, a lonely mental journey through and about meanings and contexts. It can sometimes lead to organization of knowledge through the discovery of structuring principles (Bialystok, 1991, p. 70). But this aspect of negotiation is rarely available for direct observation and analysis.

The other two aspects of negotiation—*interaction* and *interpretation*—are largely interpersonal, involving joint exploration of meaning between participants in a communicative event. Meaning cannot be conveyed entirely by surface level syntactic and semantic specifications. It has to be derived through negotiation between interlocutors. Unlike introspection, the interactional and interpretational aspects of negotiation are indeed available for observation and analysis.

Interaction

Research carried out by Long (1981), Pica (1987) and others reveals that L2 learners need to be provided with opportunities for negotiated interaction in order to help them develop language knowledge and ability. Negotiated interaction entails the learner's active involvement in clarification, confirmation, comprehension checks, requests, repairing, reacting, and turn-taking. It also means that the learner should be given the freedom and encouragement to initiate talk, not just to react and respond to it.

Several experimental studies reveal the importance of negotiated interaction. We now know that modified input and modified interaction together accelerate the rate of L2 development (Long, 1981). We also know that learners who maintained high levels of interaction in the L2 progressed at a faster rate than learners who interacted little in the classroom (Seliger, 1983) and that learners gain opportunities to develop their productive capacity in the L2 if demands are placed on them to manipulate their current IL system so that they can make their initially unclear messages become meaningful to their interlocutors (Swain, 1985). These results have been replicated by Pica and her colleagues (Pica, 1987, 1991; Pica, Young, & Doughty 1987) who report that what enables learners to move beyond their current IL receptive and expressive capacities are opportunities to modify and restructure their interaction with their interlocutor until mutual comprehension is reached. These studies lend credence to an earlier claim by Allwright (1984, p. 9) that "the importance of interaction is not simply that it creates learning opportunities, it is that it constitutes learning itself."

Interpretation

Closely associated with the opportunity to interact is the ability to interpret target language utterances as intended. It constitutes the language knowledge/ability to differentiate what is said from what is meant. Inability to do so results in pragmatic failure (Thomas, 1983). The L2 learner's interpretive ability entails an understanding of pragmatic rules such as those associated with the Hymesian concept of communicative appropriacy and the Gricean maxims of conversational implicature.

Interpretive procedures have implications for L2 development, for, as Widdowson (1983, p. 106) points out, they are "required to draw systemic knowledge into the immediate executive level of schemata and to relate these schemata to actual instances." The L2 learner encountering TL instances has to learn to deal with several possibilities:

> (a) Utterances may convey more than their literal meaning. *It's cold in here,* when spoken in certain contexts may convey the meaning of, *Would you mind closing the window?*
>
> (b) Utterances may not convey their literal meaning. *How are you* in English is often not answered at all. When it is, the speaker does not expect to learn about the hearer's ailments.
>
> (c) Utterances may convey the intended meaning only if they are accompanied by certain specifications: In American English, as foreign students have found to their chagrin, *drop in anytime* is not a genuine invitation unless clearly followed by the mention of time and place.

As these examples show, interpretation of intended meaning becomes critical, not because the notions of request, phatic communion, or invitation are unfamiliar to L2 learners, but because these familiar notions can have linguistic realizations in L2 that are very different from those in L1. In addition, learners need to be aware that lines of interpretation are likely to diverge based on cultural background (Gumperz, 1982) as well as subcultural levels of ethnic heritage, class, geographic region, age, and gender (Tannen, 1992; Kramsch, 1993).

For a realization of the full potential of negotiation factors—particularly a positive correlation with other intake factors—the individual factor of anxiety and the affective factors of attitude and motivation (see below) may be required. Aston (1986), for instance, found that interactive classroom tasks designed to promote negotiation may indeed fail to do so if they produce tension and anxiety in the learner. Thus, in conjunction with other relevant intake factors,

negotiation factors provide ample opportunities for L2 learners to pay particular attention to new features of the linguistic input that are being currently learned, thereby contributing to activating other intake factors and intake processes.

Tactical Factors

Tactical factors refers to an important aspect of L2 development: the learner's awareness of, and practical ability to use, appropriate tactics for effective learning of the TL, and efficient use of the limited repertoire developed so far. In the L2 literature, such tactics are discussed under the general rubrics of learning strategies and communication strategies.

Learning Strategies

Learning strategies are operations and routines used by the learner to facilitate the obtaining, storage, retrieval, and use of information (Rubin, 1975). They are also "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (Oxford, 1990, p. 8). The term *learning strategies*, then, refers to what learners knowingly or unknowingly do to regulate their learning.

It was only during the 70s that researchers began to study systematically the explicit and implicit efforts learners make to learn their L2 (Rubin, 1975; Naiman, Frohlich, Stern, & Todesco, 1978). Major typologies proposed so far (Rubin, 1975; O'Malley & Chamot, 1990; Oxford, 1990; Wenden, 1991) classify learning strategies into at least three broad categories: *metacognitive*, *cognitive*, and *social/affective*. *Metacognitive strategies* refers to higher order executive strategies such as thinking about the learning process, planning for and monitoring learning as it takes place, and self-evaluation of learning after the learning activity. *Cognitive strategies* refers to specific steps such as summarizing, deducing, transferring, and elaborating. *Social/affective strategies* refers to interpersonal strategies including cooperative learning, peer group discussion, and so on.

Research cited above shows that there are many individual ways of learning a language successfully and that different learners will approach language learning differently. However, it has often been found that more effective learners use a greater variety of strategies and use them in ways appropriate to the language learning task and that less effective learners not only have fewer strategy types in their repertoire but also frequently use strategies that are inappropriate to the task (O'Malley & Chamot, 1990). One of the primary objectives of research on learning strategies is to make the intuitive knowledge possessed by

good language learners more explicit and systematic, so that such a knowledge can be used for strategy training to improve the language learning abilities of other learners.

Communication Strategies

Communication strategies are "potentially conscious plans for solving what to an individual presents itself as a problem in reaching a particular communicative goal" (Faerch & Kasper, 1980, p. 81). The earliest taxonomy of communication strategies is the one proposed by Tarone (1977). It has three broad categories: *paraphrase*, *borrowing*, and *avoidance*. *Paraphrase* includes approximation, word coinage, and circumlocution. *Borrowing* includes literal translation, language switch, appeal for assistance, and mime. *Avoidance* includes topic avoidance and message abandonment.

The Tarone taxonomy in one way or another relates to interlingual, intralingual, or paralingual features. In other words, it is a product-oriented, surface-structure framework which conflates the distinction between linguistic realizations and mental processes. Bialystok and Kellerman (1987) point this out and consider in detail an example given by Tarone for the communication strategy of word coinage, *airball* for *balloon*, and use it to explain the flaw in the product-based taxonomical approach: "If learner A describes a balloon as a *ball with air* and learner B says *an airball*, then a traditional product-oriented taxonomy would call the first utterance *circumlocution*, and the second *word coinage*" (p. 164) even though both *a ball with air* and *airball* refer to identical sets of criterial attributes. Bialystok and Kellerman stress the need to go beyond IL production and to differentiate surface level communication strategies from deep level psychological processes.

Accordingly, Bialystok and Kellerman (1987) and Bialystok (1990) suggest that the strategic behavior of learners can be classified into *linguistic* and *conceptual* strategies. The linguistic strategy refers to the use of features and structures from another language (usually L1), and the conceptual strategy refers to the manipulation of the intended concept. They further divide conceptual strategy into two possible approaches: *holistic* and *analytic*. The holistic approach involves using a similar referent, as in *stove* for *microwave*. The analytic approach involves selecting criterial properties of the referent, as in *a machine that cooks and defrosts very fast by means of waves*, for *microwave*. While scholars differ on the relative explanatory power of various taxonomies, there is near unanimity concerning the facilitating role played by tactical factors in L2 development. Tactical factors can help learners pay attention to potentially useful linguistic input, thereby contributing to its recognizability (see **Input**, above).

Affective Factors

The individual learner's disposition to learn has always been recognized as an important variable in L2 development. The term *affective factors* refers to two closely connected variables that characterize learner disposition: attitudes and motivation. L2 researchers initially studied the two variables together, proposing a linear relationship in which attitude influenced motivation and motivation influenced L2 development (Gardner, 1985). Recent research, however, indicates the usefulness of separating them (Crookes & Schmidt, 1991).

Attitudes

Attitudes are one's evaluative responses to a person, place, thing, or event. According to social psychologists, attitudes are individually driven; that is, they are one's personal thoughts or feelings based on one's beliefs or opinions; therefore, different individuals develop different shades of attitudes towards the same stimuli. Attitudes are also socially grounded; that is, they must be experienced as related to subjects or events in the external world. To a large extent, an individual's attitudinal behavior is determined by social constructs, making it broadly predictable (Eiser, 1987).

In the context of L2 development, there are two forces which appear to shape the learner's language attitude: *socio-educational* (discussed under *Environmental Factors,* below) and *pedagogic.* From a pedagogic point of view, teachers, learners, and the learning situation can interact to trigger positive or negative attitudes in the learner. One of the reasons for differential success among L2 learners is their attitude towards learning the language in a particular situation, a positive attitude about language learning being a necessary but not sufficient condition for success (Naiman, Frohlich, Stern, & Todesco, 1978).

The teacher's objectives, activities, and attitudes also play a role in influencing the learner's attitude to language learning (Malcolm, 1987). In fact, teachers' attitudes seem to have a greater influence on L2 development than even parental or community-wide attitudes (Tucker & Lambert, 1973). Furthermore, a review of diary studies shows that learners can hold negative attitudes towards the learning situation if there is a mismatch between their curricular objectives and their teacher's (Schumann & Schumann, 1977).

Learner attitude towards speakers of the TL, and its impact on L2 development, have been widely studied, resulting in conflicting findings. Early experiments conducted by Gardner and his colleagues (see, for instance, Gardner & Lambert, 1972) showed high correlation between the learner's positive attitude towards speakers of the TL, and L2 development. Such a conclusive claim has since been questioned (Oller,

Baca & Vigil, 1977; Cooper & Fishman, 1977). Recent research, however, shows that, although L2 learners might develop a negative attitude towards the TL community for cultural or political reasons, a positive attitude towards the TL itself and its usefulness can contribute to L2 development (Berns, 1990).

Motivation

Motivation is perhaps the only intake variable that has been consistently found, in various contexts and at various levels of L2 development, to correlate positively with success. Most studies on motivation have been inspired by the distinction (Gardner & Lambert, 1972) between *integrative* and *instrumental* motivations. *Integrative motivation* refers to an interest in learning an L2 in order to integrate or at least interact with members of the TL community. *Instrumental motivation* refers to an interest in learning an L2 for functional purposes such as getting a job or passing an examination. In several studies, Gardner, Lambert, and colleagues (see Gardner, 1985, and the references cited there) reported that integrative motivation is far superior to instrumental motivation.

Studies conducted in other learning/teaching contexts (Lukmani, 1972; Chihara & Oller, 1978) failed to show the superiority of integrative motivation. In fact, a comprehensive review of motivational studies found a wide range of correlations covering all possibilities: positive, nil, negative, and ambiguous (Au, 1988). Recent studies by Gardner and his colleagues (Gardner & MacIntyre, 1991) clearly demonstrate that both integrative motivation and instrumental motivation have "consistent and meaningful effects on learning, and on behavioral indices of learning" (Gardner & MacIntyre, 1991, p. 69).

It is now fairly clear that the binary approach proposed by social psychologists does not adequately explain the perceived correlation between motivational types and L2 development. It may be beneficial to turn to cognitive psychologists who have suggested three types of motivation: *intrinsic*, *extrinsic*, and *achievement*.[4]

Intrinsic motivation is the desire to engage in activities characterized by enjoyment (Csikszentmihalyi, 1975; Deci, 1975; Deci & Ryan, 1985). There is no apparent reward except the experience of enjoying the activity itself. According to Csikszentmihalyi (1975), true enjoyment accompanies the experience of what he calls *flow*, that peculiar, dynamic, holistic sensation of total involvement with the activity itself. Thus, intrinsically motivated activities are ends in themselves rather than means to an end. Individuals seek out and engage in intrinsically motivated activities in order to feel competent and self-determining. Like basic human drives, intrinsic needs are

innate to the human organism and function as an important energizer of behavior.

Unlike intrinsic motivation, extrinsic motivation can be triggered only by external cues, which include gaining and maintaining peer, sibling, or adult approval; avoiding peer, sibling, or adult disapproval; and gaining or losing specific tangible rewards. It is conditioned by the practical considerations of life with all its attendant sense of struggle, success, or failure. Thus, extrinsic motivation is associated with lower levels of self-esteem and higher levels of anxiety, compared to intrinsic motivation.

Achievement motivation refers to motivation and commitment to excel. It is involved whenever there is competition with internal or external standards of excellence. It is a specific motive that propels one to utilize one's fullest potential (McClelland, Atkinson, Clark, & Lowell, 1953; Deci, 1975; Deci & Ryan, 1985).

It may be hypothesized that all threè types of motivation will influence L2 development in different degrees, depending on individual dispositions and different socio-educational contexts. To be primarily motivated for intrinsic reasons, learners have to get involved in continual cycles of seeking language learning opportunities and conquering optimal challenges in order to feel competent and self-determining. They have to let their natural curiosity and interest energize their language learning endeavor and help them overcome even adverse pedagogic and environmental limitations. To be primarily motivated for achievement considerations, the learners have to strive to reach internally induced or externally imposed standards of excellence, in a spirit of competition and triumph. It appears reasonable to assume that a vast majority of L2 learners are primarily motivated for extrinsic reasons. In fact, extrinsic motivation accounts for most of what has been reported in terms of integrative and instrumental motivation (van Lier, 1991).

The relationship between intrinsic, extrinsic, and achievement motivations is yet undetermined. The general trend of experimental studies has been to suggest that the relationship is essentially unstable and nonlinear, and that, over time, several intake factors, particularly individual, affective, and environmental factors, contribute to shape the relationship. Thus, the interplay of input, intake factors, and intake processes appears to influence the role of affective factors in L2 development.

Knowledge Factors

Knowledge factors refers to language knowledge and metalanguage knowledge. All adult L2 learners exposed to formal language education in their L1 inevitably bring with them not only their L1 knowledge/ability but also their own perceptions and expectations about

language, language learning, and language teaching. Both language knowledge and metalanguage knowledge are implicitly or explicitly present all the time in the L2 learner's mind, and hence play a crucial role in L2 development.

Language Knowledge

Language knowledge represents L2 learners' knowledge of and ability in the language system(s) already known to them, and the developing knowledge/ability of the TL they are currently learning. All adult L2 learners minimally possess L1 knowledge/ability by virtue of their experience and also by virtue of being members of their speech community. Empirical evidence shows that the L2 user does not "effectively switch off the L1 while processing the L2, but has it constantly available" (Cook, 1992, p. 571).

The influence and use of language knowledge can be a facilitating or a constraining factor in L2 development. As Corder (1983, p. 91) suggests, prior language knowledge "created and remembered from the learner's own linguistic development" may very well provide the starting point (*initial hypothesis*) of the L2 developmental continuum. Prior knowledge may also impose a set of constraints on "the domains from which to select hypotheses about the new data one is attending to" (Schachter, 1983, p. 104). In addition, through the intake process of transfer, language knowledge intersects with other intake factors and intake processes across various phases of L2 development (see *Grammaticalization*, below, for details on language transfer as an intake process).

Metalanguage Knowledge

Metalanguage knowledge is considered to be an important facilitator of L2 development (Gass, 1983; Donato & Adair-Hauck, 1992; Green & Hecht, 1992). It is "an individual's ability to match, intuitively, spoken and written utterances with his/her knowledge of a language" (Masny & d'Anglejan, 1985, p. 176). It encompasses learners' knowledge/ability not only to analyze their own language but also to make comparisons between their LI and L2, between LI and other languages previously learned, and between L2 and other languages previously learned.

There seems to be a strong relationship between language experience and metalanguage knowledge. Empirical studies reveal that prior language experience helps L2 learners develop an intuitive "feel" for the TL (Gass, 1983; Donato & Adair-Hauck, 1992; Green & Hecht, 1992). L2 learners have been shown to be able "to produce a correct correction when they have an incorrect explicit rule or no explicit rule at all,"

thereby demonstrating the presence of L2 intuitions (Green & Hecht, 1992, p. 176). Extending the role of metalanguage knowledge, Cook (1992) has recently proposed the concept of *multicompetence* to describe "the compound state of a mind with two grammars" (p. 558) in contrast to *monocompetence*, the state of mind with only one grammar. According to him, the multicompetent individual approaches language differently in terms of metalinguistic awareness. Cook hypothesizes that such a heightened metalinguistic awareness may impact other aspects of cognition, thereby shaping the cognitive processes of L2 development. There is thus both theoretical and empirical evidence to support the view that the knowledge factor plays an important role in L2 development. In fact, knowledge as an intake factor is much more than language and metalanguage knowledge/abilities put together. Perhaps a more apt term is *prior text*, as used by Becker (1983). In relation to L2 experience, the role of prior text is to help the learner characterize the present in the past and "to make any new utterance reverberate with past ones, in unpredictable directions" (Becker, 1983, p. 218).

Environmental Factors

Environmental factors refers to social, cultural, political, economic, educational, and technological milieus in which L2 learning and teaching take place. The impact of these factors on L2 development has not been fully explored; however, there are indications that L2 development is highly responsive to social and educational contexts.

Social Context

Social context refers to a range of language learning environments such as the home, the neighborhood, the classroom, and the society at large. Any serious attempt to study L2 development necessarily entails the study of social context as an important variable (Heath, 1983; Beebe, 1985; Breen, 1985; Sridhar & Sridhar, 1986; Wolfson, 1989; Wong-Fillmore, 1989; Berns, 1990; Kramsch, 1993). In fact, Beebe (1985) argues that the learner's choice of what input becomes intake is highly affected by social and situational contexts. Additionally, social context is critical because it shapes various learning/teaching issues such as the motivation for L2 learning, the goal of L2 learning, the functions L2 is expected to perform in the community, the availability of input to the learner, the variation in the input, and the norms of proficiency acceptable to that particular speech community.

Specific social settings such as the neighborhood and the classroom in which learners come into contact with the new language have also been found to influence L2 development. Studies conducted by Wong-Fillmore (1989) reveal that social settings create and shape

opportunities for both learners and competent speakers of the L2 to communicate with each other, thereby maximizing learning potential. A recent study by Donato and Adair-Hauck (1992) concludes that the social and discursive context in which instructional intervention is delivered plays a crucial role in facilitating L2 development in the classroom.

The social context also shapes the role of the TL in a particular speech community and the nature of the linguistic input available for learners. Comparing the sociolinguistic profiles of English language learning and use in India, West Germany, and Japan, Berns (1990) illustrates how these three different social contexts contribute to the emergence of various communicative competencies and functions in these countries, thereby influencing L2 development and use in significantly different ways. In these and similar contexts, the TL plays a role that is complementary or supplementary to local/regional language(s). These competencies and functions invariably determine the nature and quality of input that is available to the learner. Most often, the learner is not exposed to the full range of the TL in all its complexity that one would expect in a context where it is used as the primary vehicle of communication.

Educational Context

Closely related to social context is educational context. Studies on educational contexts grounded in educational psychology emphasize the inseparability and reciprocal influence of educational institutions and settings in which learning/teaching operations are embedded (Bloome & Green, 1992). Although the educational context in which learning occurs shapes language learning abilities, its exact influence on L2 development has not been fully explored. L2 development may seem like a discrete activity, but it is actually grounded in larger educational contexts that have profound effect on learning. For instance, it is the educational context which shapes policy constraints, language planning, and most importantly, the learning opportunities available to the L2 learner. It is impossible to insulate classroom life from the dynamics of political, educational, and societal institutions (Kachru, 1990; Tollefson, 1991).

To sum up this section, all the intake factors outlined above—individual, negotiation, tactical, affective, knowledge, and environmental—appear to interact with each other in as yet undetermined ways. They play a significant role in triggering and maximizing the operational effectiveness of intake processes, to which we turn now.

Intake Processes

Intake processes are internal operations that at once mediate between, and respond to, input and intake factors. They consist of operations which are specific to language learning as well as those which are required for general problem-solving. They are linguo-cognitive in nature; that is, they are either primarily linguistic with an added cognitive dimension, or primarily cognitive with an added linguistic dimension. As procedures and operations that are internal to the learner, intake processes and their interrelationships remain the most vital and least understood link in the input-intake-output chain. In the rest of this section, I outline each of these processes.

Linguistic Processes

Intake processes that are primarily linguistic in nature include *grammaticalization*, a process that involves linguistic structures or functions common to most natural languages, and *language transfer*, a process that involves the interplay between earlier and later learned language systems.

Grammaticalization

The process of grammaticalization refers primarily to the role played by Universal Grammar (UG), an element of biologically endowed genetic principles common to the human species. A review of research shows two broad claims with regard to the role played by UG principles in L2 development: (a) all of them operate in L2 as they do in L1; and (b) some of them operate in L2 but not in the same way as they do in L1. Researchers such as Gass and Ard (1984); Flynn (1987); Liceras (1989); and White (1990) maintain that L2 learners have access to the same innate constraints and properties of UG as do children. They say so primarily because there are syntactic representations in child L1 and adult L2 production which cannot be induced simply from the available linguistic input. They point out that in order to form and test hypotheses about correct and incorrect language forms, learners need two kinds of evidence, positive and negative. In the context of L2 development, positive evidence comes from language input which contains well-formed utterances, and negative evidence comes from feedback to learners in the form of direct or indirect correction. Children get hardly any negative evidence, and the negative evidence that L2 learners get is certainly inadequate. Therefore, it is assumed that child L1 acquirers and adult L2 learners have access to the same innate universal constraints and properties.

Researchers such as Schachter (1988), Bley-Vroman (1988), Clahsen (1990), and Sharwood Smith (1991) question the preeminence of UG principles in determining L2 development. Pointing out the ways in which L2 learning differs from L1 acquisition, Bley-Vroman (1988) has proposed a Fundamental Difference Hypothesis which states that L1 knowledge and general problem-solving capacity of the adult L2 learner assume much (not all) of the role played by UG in child L1 acquisition. Extending the Bley-Vroman proposal, Clahsen (1990, pp. 150-151) offers "a particular version of the Fundamental Difference Hypothesis, according to which (a) parameterized UG principles are lost in adult L2 learners, and (b) stable UG principles are present only through the learner's first language." Based on the findings available at present, it seems reasonable to assume that some sort of UG does continue to play a role in adult L2 development, but not in the same significant way it does in child L1 acquisition.

Yet another strand of grammaticalization emphasizes the projective power that is supposed to enable the acquisition of one rule to trigger the acquisition of all the other rules that are implicationally linked to it. Gass (1979) tested the projection hypothesis by using the Relative Clause Accessibility Hierarchy proposed by Keenan and Comrie (1977); her study showed that the learners not only succeeded in improving their scores on the one relative clause structure that was the focus of instruction, but also on all the positions higher in the hierarchy. Eckman, Bell, and Nelson (1988) replicated the Gass study with a more rigorous research design and found that their learners also generalized instruction to other related structures when they were taught only one particular structure. Similar findings about the influence of typological universals on L2 development have been reported by Givon (1984) with regard to topic continuity hierarchy, and by Zobl (1985) with regard to the human > nonhuman markedness scale. These experimental studies reveal that learners learn not only those features that have been taught but also other features that are implicationally associated with them.

Language Transfer

Language transfer encompasses a whole range of behaviors, processes, and constraints, each of which has to do with the influence and use of the developing TL as well as the language systems already known to the L2 learner (Selinker, 1992). Drawing insights from a series of empirical studies (see the volumes edited by Gass & Selinker, 1983; Davies, Criper, & Howatt, 1984; Kellerman & Sharwood Smith, 1986), we can now conceptualize language transfer not as a mechanical transfer of L1 structure, but as a complex linguistic/cognitive process involving many factors which operate at syntactic, semantic, phonological, and discourse

levels. We learn that differences between languages in contact do not necessarily cause difficulties and that similarities do not necessarily facilitate development. We also learn that there are psychotypological constraints which result not from surface level similarities and dissimilarities but from the learner's perceptions of language distance, language specificity, and language universality (Kellerman, 1983).

In a comprehensive review of the literature on language transfer, Selinker (1992) concludes that interlingual identifications from L1 to L2 input are essential to the formation of IL and that transfer effects do not occur in an absolute all-or-nothing fashion. That is, learners do not transfer entire phonological, morphological, or syntactic systems of L1; instead, they select what to transfer and what not to transfer. Their selection process is presumably facilitated or constrained by transfer as well as by the process of grammaticalization.

Cognitive Processes

Intake processes that are primarily cognitive include *inferencing*, *structuring* and *restructuring*. These processes guide what learners have to do to develop their L2: to infer from the available/recognized input data the linguistic system of the TL, to structure appropriate mental representations of the TL system, and to restructure the developing IL system.

Inferencing

The intake process of *inferencing* involves making informed guesses to derive working hypotheses about various aspects of the TL system by using all available—and at times inconclusive—linguistic evidence which includes intralingual and interlingual cues, as well as non-linguistic evidence which includes the learner's knowledge of the world. Working hypotheses so derived may lead to interim conclusions which are tested against new evidence and subsequently rejected or refined. Inferencing thus may entail framing new insights or reframing what is already vaguely or partially known.

Learners may have at their disposal three types of inferencing attributable to at least three sources: *implicit knowledge*, *other knowledge*, and *context* (Bialystok, 1983). *Implicit knowledge* refers to information the learners intuit about the TL, even though they cannot articulate that information in the form of rules or principles. *Other knowledge* refers to the learners' knowledge about the TL, their L1, and their knowledge of the world. *Context* includes both linguistic and physical aspects of a situation which provide input. Inferencing is successful only to the degree that the learners are able to make connections between these three sources of knowledge on the one hand, and between them

and the input data on the other. Inferencing can be expected to vary from learner to learner because it reflects individual cognitive capabilities involving connections made by learners themselves, and not connections inherently found in the input data.

Structuring

The intake process of *structuring* combines elements of analysis and control proposed by Bialystok (1988, 1990). The formation of mental representations of the TL and their evolution in the course of IL development may be called *structuring*. The process helps learners construct and organize the symbolic representational system of the TL by gradually making explicit the implicit knowledge that shape their IL performance. Structuring also guides the gradual progress learners make, from unanalyzed knowledge consisting of prefabricated patterns and memorized routines, to analyzed knowledge consisting of propositions in which the relationship between formal and functional properties of the TL become increasingly apparent to the learners.

Compared to inferencing, structuring gives learners not only a greater control over the properties and principles governing the TL system, but also a greater ability to articulate them. It helps them pay selective attention to relevant and appropriate input data in order to tease out specific language problems. Structuring can also regulate the flow of information between short-term and long-term memory systems taking the responsibility for differential applicability of interim knowledge to various situations before interim knowledge gets fully established. The difference between inferenced knowledge/ability and structured knowledge/ability may contribute to the distinction Chaudron (1983, pp. 438-439) makes between *preliminary intake* and *final intake*. The former relates to "perception and comprehension of forms" and the latter to "the incorporation of the forms in the learner's grammar." Although inferenced knowledge/ability and structured knowledge/ability are partially independent and partially interacting dimensions of intake processes, they constitute two ends of a continuum.

Restructuring

The idea of *restructuring* as an intake process is derived from the work of Cheng (1985) in cognitive psychology and applied with some modification to L2 development by McLaughlin and his colleagues (McLeod & McLaughlin, 1986; McLaughlin, 1987, 1990). Restructuring can be traced to the Piagetian approach, which maintains that cognitive development is characterized by fundamental qualitative change when a new internal organization is imposed for interpreting new information. In other words, restructuring denotes neither an

incremental change in the structure already in place nor a slight modification of it, but the addition of a new structure to allow for a new interpretation. It marks a strategy shift that coordinates, integrates, and reorganizes task components, resulting in more efficient intake processing. It can operate at phonological, morphological, syntactic, semantic, and discourse levels (McLaughlin, 1990).

While most aspects of inferencing and structuring account for the reasons why intake processing requires selective attention and an extended time period for the formation of mental representations of the TL system, restructuring as an intake process accounts for discontinuities in L2 development. It has been frequently observed that while some learning occurs continuously and gradually, as is true of the development of automaticity through practice, some learning occurs in discontinuous fashion, through restructuring (McLeod & McLaughlin 1986). Restructuring is mostly a sudden, abstract, insight-forming phenomenon happening quickly and incidentally, taking very little processing time and energy.

To sum up, intake processes are primarily linguistic and cognitive mechanisms. Linguistic mechanisms of grammaticalization and language transfer, and cognitive mechanisms of inferencing, structuring and restructuring work together in as yet undetermined ways to facilitate or constrain L2 development. These intake processes seem to operate at various points on the implicit-explicit continuum, triggering incidental learning at some times and intentional learning at other times. In combination with various intake factors, these processes help learners to synthesize the developing knowledge into grammar and internalize it so as to effectively and efficiently access it in appropriate contexts.

Output

Output refers to the corpus of utterances which learners actually produce orally or in writing. In addition to well-formed utterances that may have already been structured and/or restructured, the learner output will contain (as discussed under **Input**, above) deviant utterances which cannot be traced to any of the three major sources of input, since they are the result of an interplay between intake factors and intake processes.

An Interactive Framework of Intake Processes

Having briefly discussed various aspects of input, intake, intake factors, intake processes, and output, I shall now attempt to pull these constructs together and propose what I call an interactive framework of intake processes. First, it seems reasonable to posit two criteria that any framework of intake processes must necessarily satisfy: The framework

must be capable of including all the intake factors known to play a role in intake processes; and it must reflect the interactive and parallel nature of intake processes.

The first criterion is explicit in the L2 literature. As the discussion (under **Intake Factors,** above) amply shows, there are several learner-internal and learner-external intake factors of varying importance that, separately or in combination, facilitate or constrain L2 development. The issue facing current investigations is not whether any of the intake factors promote L2 development but how many, to what extent, in what combination and in what context.

The second criterion emerges from current theories in cognitive psychology and information processing (Anderson, 1983; McClelland, Rumelhart, & the PDP Research Group, 1986). It has been reported that cognitive processing goes on simultaneously in many areas and at many different levels. Language learning might entail a non-linear, parallel, interactive process rather than a linear, serial, additive process. It was earlier believed that learners internalize the TL system primarily by using either a top-down processing, a knowledge-governed system characterized by a step-by-step progression in which output from one level acts as input for the next; or a bottom-up processing, an input-governed system characterized by a serial movement of information from the lower to the higher levels. It is now believed that language learning is governed by interactive processing in which multiple operations occur simultaneously at multiple levels, drawing evidence from multiple sources.

Given the two criteria mentioned above, the proposed interactive framework of intake processing assumes that input, intake factors, and intake processes play a coordinated role in constraining or facilitating L2 development. Language processing is considered essentially interactive, involving intake factors and intake processes which operate in parallel and simultaneous ways, shaping and being shaped by one another.

The proposed interactive framework consists of input, intake factors, a central processing unit (CPU), and output. The CPU includes linguistic processes of grammaticalization and language transfer, and cognitive processes of inferencing, structuring, and restructuring.

As *Figure 5* indicates, intake processing is activated when input information enters the CPU either directly or through one or more intake factors. One can speculate that the basic properties of naturally endowed formal universals, as well as other linguistic/cognitive knowledge/abilities that adult learners bring to bear on L2 development initiate the process of language construction. At this early stage, intake processing appears to operate at several layers, some of which may depend heavily on temporary, limited-capacity, working-memory systems

which in turn involve to a large degree prefabricated routines and idiomatic expressions.

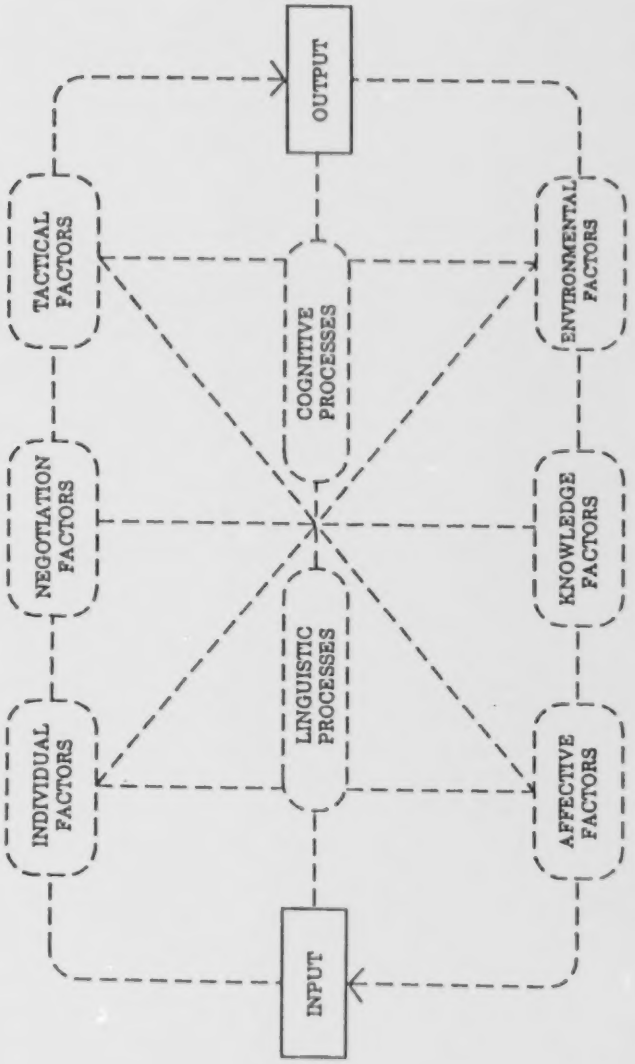


FIGURE 5
An Interactive Framework of Intake Processes

605

605

At this stage, an important task of the CPU is to reduce the pressure on working memory systems by coding the incoming information according to certain organizational schemas. Such a coding, which is probably a precursor to fully established mental representations, is assisted by the intake process of inferencing. Inferencing helps learners derive working hypotheses about syntactic, semantic, and pragmatic aspects of the TL. Positive evidence in the form of additional input data, and negative evidence in the form of feedback from competent speakers of the TL, make the learners reject or refine working hypotheses.

If the process of inferencing constitutes a designing of working hypotheses, the process of structuring constitutes a devising of mental representations. As we learn from schema theory, the faster the testing and refinement of working hypotheses, the swifter the formation of mental representations and the greater the chances of limited-capacity, working-memory systems being purged and replaced by permanent long-term memory schemas. Memory schemas are responsible for storing incoming information, retrieving previously stored information, and pattern-matching mental representations (McClelland, Rumelhart, & the PDP Research Group, 1986). This transition from working memory systems to permanent memory schemas is critical because, as we learn from schema theorists, language use requires that linguistic units such as phonemes, morphemes, words, phrases, syntactic patterns, and other discourse units be abstracted and stored in the form of memory schemas.

Repeated cycles of rejection and refinement of working hypotheses, and the construction of memory schemas mediated by intake processes, particularly by the process of structuring, result in the establishment of mental representations of the TL, thereby considerably increasing learners' ability to gain control over the properties and principles of the TL system. Any remaining gap in the establishment of mental representations is taken care of either by further opportunities for intentional corrective learning or by the activation of the process of restructuring. Restructuring represents a process of quick insight formation that could result in incidental learning whereby complex and hitherto unclear language problems are teased out, paving the way for accurate decisions about the TL system.

Each of the intake processes is constrained not merely by the availability/recognizability of linguistic input and the interplay of intake factors but also by the role played by learner output. The arrows connecting input and output (*Figure 5*) suggest that the latter is not a terminal point; it is rather a part in a cycle serving as an important source of input data for the learner, thereby affecting the course of L2 development (Swain, 1985; Schmidt & Frota, 1986).

In explaining intake processing, the interactive framework proposed here incorporates several aspects of parallel distributed processing[5] at both micro and macro levels. At the micro level, intake

processing is considered to involve a large number of parallel, simultaneous, and interacting processes such as perception, syntactic parsing, and semantic interpretation, and the selection of whatever information is relevant and useful, be it phonological, syntactic, semantic, or pragmatic. The development of some syntactic rules, for example, often depends on the development of a rule in some other domain, say a phonological or lexical rule, or vice versa (Ard & Gass, 1987; Klein, 1990). Following the connectionist perspective, the intake processing network is seen as a continual strengthening or weakening of interconnections in response to examples encountered in the input data, and experience in using the developing system.

At the macro level, the framework posits a criss-cross interplay among intake factors on the one hand, and between them and intake processes on the other. Most of the intake factors appear to interweave and interact with each other in a synergic relationship where the whole is greater than the sum of the parts. To draw an analogy, the intake factors function much like the subsystems of our ecological system in the sense that each subsystem operates to influence, and to be influenced by, the other. How the learner seeks, recognizes, attends to, and controls the input data depends to a large extent on the synergy of intake factors.

The interactive framework also suggests that the linguistic input is not processed linearly, proceeding step-by-step from one intake factor through another, or from one intake process through another. Instead, the entire operation is seen as interactive and parallel, responding simultaneously to all available factors and processes at a given point of time. In other words, none of the intake factors by itself seems to be a *prerequisite* for another to be activated, but all are considered *co-requisites*. The processing of input data is never consistent; it varies according to varying degrees of influence brought to bear on it by unstable and as yet unknown configurations of intake factors and intake processes. Different intake factors and processes take on different statuses in different "ecological conditions," thereby significantly affecting learners' working hypotheses about the TL and their strategies for learning and using it. The configuration also varies widely within an individual learner at different times and situations of learning, and also among learners, thereby accounting for wide variations in degree of attainment.

Conclusion

This paper explored the concepts of intake, intake factors, and intake processes in order to interpret the factors and processes facilitating adult L2 development in formal contexts. It has been argued here that any framework of intake processing must be capable of including multiple intake factors known to play a role in L2

development, and that it must reflect the interactive, parallel, and simultaneous nature of intake processes. Accordingly, this paper attempted to design an interactive framework by synthesizing theoretical and empirical insights derived from interrelated disciplines such as second language acquisition, cognitive psychology, information processing, schema theory, and parallel distributed processing.

In addition to input and output, the interactive framework of intake processes presented here consists of a cluster of intake factors (Individual, Negotiation, Tactical, Affective, Knowledge and Environmental factors) and intake processes (linguistic processes of grammaticalization and language transfer, and cognitive processes of inferencing, structuring, and restructuring). Interweaving and interacting in a synergic relationship, each intake factor shapes and is shaped by the other. The interactive nature of intake factors and intake processes suggests that input can be successfully converted into intake if, and only if, the intake factors and processes are optimally favorable and if the consistent absence of one or a combination of these constructs may result in partial learning.

The interactive framework presented here casts doubts about the nature and scope of current research in L2 development. For the past 25 years, we have been focusing mostly upon narrowly circumscribed research problems within each intake variable, accumulating an impressive array of unrelated and unrelatable findings which, by the very nature of investigation, can allow only a limited and limiting view of L2 development. If, as this paper emphasizes, several intake factors facilitate the course of L2 development; if these factors are mutually dependent; if they shape and are shaped by each other; and if they are constantly acted upon by intake processes which are interactive, parallel, and simultaneous, then it is imperative that we reframe our research agenda by focusing on the synergic relationships between and within intake factors and processes in order to find, as anthropologist Gregory Bateson would say, "the pattern that connects."

Notes

1. To avoid the conceptual connotations attached to *acquisition* and *learning*, I use the theory-neutral term *development*, which indicates that L2 knowledge cannot be easily dichotomized as acquired/learned or implicit/explicit.

2. Unlike the proponents of intake as product or intake as process, Boulouffe (1986, p. 258) treats intake as a dual construct having both product and process components: "Intake as a process originates in the speaker's intent, and the learner's learning strategies initiate a debate

608

between assimilation and accommodation. Intake as a product is the result of intake as a process."

3. I use *knowledge/ability* instead of *competence* and *performance*, which have been used variously in various contexts. *Knowledge* refers to the notion of knowing the language, and *ability* refers to the practice of using the language. To indicate that knowledge and ability are not clearly separable dichotomies, I use *knowledge/ability* as a compound term. For an in-depth discussion on this issue, see the 1989 thematic issue of *Applied Linguistics, 10* (2).

4. The intrinsic-extrinsic approach is not new to L2 research. Scholars have made use of this approach from time to time. Recently, Brown (1990); Crookes, and Schmidt (1991); and van Lier (1991) have either used or suggested the use of this approach.

5. The general idea of parallel, simultaneous, multiple processes suggested in the connectionist/PDP model is useful for interpreting intake processing. However, some of the other tenets of the PDP model may not be applicable to adult L2 development. For instance, the model (a) does not take into account important intake factors such as affective factors, environmental facts that may influence intake processes, and (b) starts with some form of *tabula rasa*, which would seem to disallow any innate linguistic abilities as a driving force in L2 development (see Gasser, 1990, for a detailed critique).

References

Allwright, R. L. (1984). Why don't learners learn what teachers teach? —The interaction hypothesis. In D. M. Singleton & D. G. Little (Eds.), *Language learning in formal and informal contexts* (pp. 3-18). Dublin: IRAAL.

Alpert, R. & Haber, R. (1960). Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology, 61*, 207-215.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Ard, J. & Gass, S. M. (1987). Lexical constraints on syntactic acquisition, *Studies in Second Language Acquisition, 9*, 233-251.

Aston, G. (1986). Trouble-shooting in interaction with learners. The more the merrier? *Applied Linguistics, 7*, 128-143.

Au, S. Y. (1988). A critical appraisal of Gardner's social-psychological theory of second language (L2) learning. *Language Learning, 38*, 75-100.

Bailey, K. M. (1983). Competitiveness and anxiety in adult second

language learning: Looking at and through the diary studies. In H. Seliger & M. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 67-103). Rowley, MA: Newbury House.

Beebe, L. M. (1985). Input: Choosing the right stuff. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 404-414). Rowley, MA: Newbury House.

Becker, A. L. (1983). Toward a post-structuralist view of language learning: A short essay. *Language learning, 33*, 217-220.

Berns, M. (1990). *Contexts of competence: Social and cultural considerations in communicative language teaching.* New York: Plenum.

Bialystok, E. (1983). Inferencing: Testing the "hypothesis-testing" hypothesis. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 104-123). Rowley, MA: Newbury House.

Bialystok, E. (1988). Psycholinguistic dimensions of second language proficiency. In W. Rutherford & M. Sharwood Smith (Eds.), *Grammar and second language teaching: A book of readings* (pp. 31-50). New York: Newbury House/Harper and Row.

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second language use.* Oxford: Basil Blackwell.

Bialystok, E. and Kellerman, E. (1987). Language strategies in the classroom. In B. Das (Ed.), *Communication and learning in the classroom community* (pp. 160-175). Singapore: SEAMEO Regional Language Centre.

Bley-Vroman, R. (1988). The fundamental character of foreign language learning. In W. Rutherford & M. Sharwood Smith (Eds.), *Grammar and second language teaching: A book of readings* (pp. 19-30). New York: Newbury House/Harper and Row.

Bloome, D. & Green, J. L. (1992), Educational contexts of literacy. *Annual Review of Applied Linguistics, 12*, 49-70.

Boulouffe, J. (1986). Intake as the locus of equilibration in language learning. *Language Learning, 36*, 245-276.

Breen, M. P. (1985). The social context for language learning—A neglected situation? *Studies in Second Language Acquisition, 7*, 135-158.

Brown, D. H. (1990). M&Ms for language classrooms: Another look at motivation. In *Georgetown University Round Table on Languages and Linguistics* (pp. 383-393). Washington, DC: Georgetown University Press.

Chaudron, C. (1983). Simplification of input: Topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly, 17*, 437-458.

Chaudron, C. (1985). Intake: On models and methods for discovering

learners' processing of input. *Studies in Second Language Acquisition, 7*, 1-14.

Cheng, P. W. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review, 92*, 414-423.

Chihara, T. & Oller, J. W. (1978). Attitudes and attained proficiency in EFL: A sociolinguistic study of adult Japanese speakers. *Language Learning, 28*, 55-68.

Clahsen, H. (1990). The comparative study of first and second language development. *Studies in Second Language Acquisition, 12*, 135-154.

Cook, V. J. (1992). Evidence for multicompetence. *Language Learning, 42*, 557-591.

Cook, V. J. (1993). *Linguistics and second language acquisition*. New York: St. Martins Press.

Cooper, R. L. & Fishman, J. A. (1977). A study of language attitudes. In J. A. Fishman, R. L. Cooper, & A. W. Conrad (Eds.), *The spread of English* (pp. 239-276). Rowley, MA: Newbury House.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161-170.

Corder, S. P. (1978). Language learner language. In J. C. Richards (Ed.), *Understanding second and foreign language learning: Issues and approaches* (pp. 71-93). Rowley, MA: Newbury House.

Corder, S. P. (1983). A role for the mother tongue. In S. M. Gass. & L. Selinker (Eds.), *Language transfer in language learning* (pp. 85-97). Rowley, MA: Newbury House.

Crookes, G. & Schmidt, R. W. (1991). Motivation: Reopening the research agenda. *Language Learning, 41*, 469-512.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass.

Davies, A., Criper, C. & Howatt, A. P. R. (Eds.). (1984). *Interlanguage*. Edinburgh: Edinburgh University Press.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum.

Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Donato, R., & Adair-Hauck, B. (1992). Discourse perspectives on formal instruction. *Language Awareness, 1*, 74-89.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language Two*. New York: Oxford University Press.

Eckman, F. R., Bell, L. H., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*, 1-20.

Eiser, J. R. (1987). *The expression of attitude*. New York: Springer-Verlag.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford:

Oxford University Press.
Faerch, C. & Kasper, G. (1980). Processes and strategies in foreign language learning and communication. *Interlanguage Studies Bulletin, 5*, 47-118.
Flynn, S. (1987). A parameter-setting model of L2 acquisition. Dordrecht: Reidel.
Gardner, R. C. (1985). *Social psychological and second language learning: The role of attitudes and motivation.* London: Edward Arnold.
Gardner, R. C., Day, J. B., & MacIntyre, P. D. (1992). Integrative motivation, induced anxiety, and language learning in a controlled environment. *Studies in Second Language Acquisition, 14*, 197-214.
Gardner, R. C. & Lambert, W. E. (1972). *Attitude and motivation in second language learning.* Rowley, MA: Newbury House.
Gardner, R. C. & MacIntyre, P. D. (1991). An instrumental motivation in language study: Who says it isn't effective? *Studies in Second Language Acquisition, 13*, 57-72.
Gass, S. M. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*, 327-344.
Gass, S. M. (1983). The development of L2 intuitions. *TESOL Quarterly, 17*, 273-291.
Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9*, 198-217.
Gass, S. M. & Ard, J. (1984). Second language acquisition and the ontology of language universals. In W. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 33-68). Amsterdam: John Benjamins.
Gass, S. M. & Selinker, L. (Eds.). (1983). *Language transfer in language learning.* Rowley, MA: Newbury House.
Gasser, M. (1990). Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition, 12*, 179-199.
Givon, T. (1984). Universals of discourse structure and second language acquisition. In W. Rutherford (Ed.), *Language universals and second language acquisition.* Amsterdam: John Benjamins.
Green, P. S. & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics, 13*, 168-184.
Gumperz, J. J. (1982). *Discourse strategies.* Cambridge: Cambridge University Press.
Hatch, E. (1983). Simplified input and second language acquisition. In R. Andersen (Ed.), *Pidginization and creolization as language acquisition* (pp. 64-86). Rowley, MA: Newbury House.
Heath, S. B. (1983). *Ways with words: Language and work in*

612

communities and classroom. Cambridge: Cambridge University Press.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70,* 125-132.

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of ESL. *Cognitive Psychology, 21,* 60-99.

Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes, 9,* 3-20.

Keenan, E. & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8,* 63-99.

Kellerman, E. (1983). Now you see it, now you don't. In S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112-134). Rowley, MA: Newbury House.

Kellerman, E. & Sharwood Smith, M. (Eds.). (1986). *Cross-linguistic influence in second language acquisition.* Oxford: Pergamon Press.

Kimball, M. C. & Palmer, A. S. (1978). The dialog game: A prototypical activity for providing proper intake in formal instruction. *TESOL Quarterly, 12,* 17-29.

Klein, W. (1990). A theory of language acquisition is not so easy. *Studies in Second Language Acquisition, 12,* 219-231.

Kramsch, C. (1993). *Context and culture in language teaching.* Oxford University Press.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Krashen, S. (1983). Newmark's ignorance hypothesis and current second language acquisition theory. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 135-153). Rowley, MA: Newbury House.

Lamendella, J. (1977). General principles of neurofunctional organization and their manifestation in primary and nonprimary language acquisition. *Language Learning, 27,* 155-196.

Larsen-Freeman, D. (1983). Second language acquisition: Getting the whole picture. In K. M. Bailey, M. Long, & S. Peck (Eds.), *Second language acquisition studies* (pp. 3-22). Rowley, MA: Newbury House.

Larsen-Freeman, D. & Long, M. (1991). *An introduction to second language acquisition research.* London: Longman.

Lenneberg, E. (1967). *Biological foundations of language.* New York: John Wiley.

Liceras, J. (1985). The role of intake in the determination of learners' competence. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 354-373). Rowley, MA: Newbury House.

Liceras, J. (1989). On some properties of the "pro-drop" parameter: Looking for missing subjects in non-native Spanish. In S. M. Gass. & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 109-133). Cambridge: Cambridge University Press.

Long, M. (1981). Input, interaction and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition* (pp. 259-278). Annals of the New York Academy of Sciences, *379*.

Lukmani, Y. (1972). Motivation to learn and language proficiency. *Language Learning, 22*, 261-273.

MacIntyre, P. D. & Gardner, R. C. (1989). Anxiety and second language learning: Toward a theoretical clarification. *Language Learning, 39*, 251-275.

MacIntyre, P. D. & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41*, 125-159.

Madsen, H. S., Brown, B. L., & Jones, R. L. (1991). Evaluating student attitudes toward second-language tests. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 65-86). Englewood Cliffs, NJ: Prentice-Hall.

Malcolm, I. G. (1987). Continuities in communicative patterns in cross-cultural classrooms. In B. Das (Ed.), *Communication and learning in the classroom community* (pp. 37-63). Singapore: SEAMEO Regional Language Centre.

Masny, D. & d'Anglejan, A. (1985). Language, cognition and second language grammaticality judgments. *Journal of Psycholinguistic Research, 14*, 175-197.

McClelland, D. C., Atkinson, J. W., Clark, R. W., & Lowell, E. (1953). *The achievement motive.* New York: Appleton-Century-Crofts.

McClelland, J. L., Rumelhart, D. E., & the PDP Research Group (Eds.). (1986). *Parallel distributed processing. Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models.* Cambridge, MA: MIT Press.

McLaughlin, B. (1987). *Theories of second language learning.* London: Edward Arnold.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11*, 113-128.

McLeod, B. & McLaughlin, B. (1986). Restructuring or automatization? Reading in a second language. *Language Learning, 36*, 109-127.

Naiman, N., Frolich, M., Stern, H. H., & Todesco, A. (1978). The good language learner. *Research in Education (Series #7)*. Toronto: Ontario Institute of Studies in Education.

Oller, J., Baca, L., & Vigil, F. (1977). Attitude and attained proficiency in ESL: A sociolinguistic study of native speakers of Chinese in the United States. *Language Learning, 27*, 1-26.

O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

Oxford, R. (1990). *Language learning strategies: What every teacher should know.* New York: Newbury House/Harper and Row.

Pica, T. (1987). Second language acquisition, social interaction, and the classroom. *Applied Linguistics, 8*, 3-21.

Pica, T. (1992). The textual outcomes of native speaker-nonnative speaker negotiations: What do they reveal about second language learning? In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 198-237). Lexington, MA: D. C. Heath.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21*, 737-758.

Rubin, J. (1975). What the good language learner can teach us. *TESOL Quarterly, 9*, 41-51.

Schachter, J. (1983). A new account of language transfer. In S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 98-111). Rowley, MA: Newbury House.

Schachter, J. (1988). Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics, 9*, 219-135.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.

Schumann, F. & Schuman, J. (1977). Diary of a language learner: An introspective study of second language learning. In H. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL '77* (pp. 241-249). Washington, DC: TESOL.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* New York: Newbury House/Harper and Row.

Seliger, H. W. (1984). Processing universals in second language

acquisition. In In F. R. Eckman, L. H. Bell, & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 36-47). Rowley, MA: Newbury House.

Selinker, L. (1992). *Rediscovering interlanguage*. London: Longman.

Sharwood Smith, M. (1985). From input into intake: On argumentation in second language acquisition. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 394-404). Rowley, MA: Newbury House.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*, 118-132.

Singleton, D. (1989). *Language acquisition: The age factor*. Clevedon: Multilingual Matters.

Snow, C. (1983). Age differences in second language acquisition: Research findings and folk psychology. In K. M. Bailey, M. Long, & S. Peck (Eds.), *Second language acquisition studies* (pp. 141-150). Rowley, MA: Newbury House.

Spolsky, B. (1989). *Conditions for second language learning*. Oxford: Oxford University Press.

Sridhar, K. K. & Sridhar, N. (1986). Bridging the paradigm gap: Second language acquisition theory and indigenized varieties of English. *World Englishes, 5*, 3-14.

Stern, H. H. (1984). *Fundamental concepts of language teaching*. Oxford: Oxford University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Tannen, D. (1992). Rethinking power and solidarity in gender and dominance. In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 135-147) Lexington, MA: D. C. Heath.

Tarone, E. (1977). Conscious communication strategies in interlanguage. In D. H. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL '77* (pp. 194-203). Washington, DC: TESOL.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*, 91-112.

Tobias, S. (1986). Anxiety and cognitive processing of instruction. In R. Schwarzer (Ed.), *Self-related cognition in anxiety and motivation* (pp. 35-54). Hillsdale, NJ: Erlbaum.

Tollefson, J. W. (1991). *Planning language, planning inequality*. London: Longman.

Tucker, G. & Lambert, W. (1973). Sociocultural aspects of language study. In J. Oller & J. Richards (Eds.), *Focus on the learner*. Rowley, MA: Newbury House.

616

van Lier, L. (1991). Inside the classroom: Learning processes and teaching procedures. *Applied Language Learning, 2*, 29-69.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

VanPatten, B., & Cadierno, T. (1993). Input processing and second language acquisition: A role for instruction. *Modern Language Journal, 77*, 45-57.

Wenden, A. (1991). *Learner strategies for learner autonomy*. London: Prentice-Hall International.

White, L. (1990). Second language acquisition and universal and grammar. *Studies in Second Language Acquisition, 12*, 121-133.

Widdowson, H. G. (1983). *Learning purposes and language use*. Oxford: Oxford University Press.

Wolfson, N. (1989). *Perspectives: Sociolinguistics and TESOL*. New York: Newbury House/Harper and Row.

Wong-Fillmore, L. (1989). Language learning in social context: The view from research in second language learning. In R. Dietrich & C. Graumann (Eds.), *Language processing in social context* (pp. 277-302). Amsterdam: Elsevier.

Zobl, H. (1985). Grammar in search of input and intake. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 329-344). Rowley, MA: Newbury House.

Author

B. KUMARAVADIVELU, Associate Professor, Department of Linguistics and Language Development, One Washington Square, San Jose State University, San Jose, CA 95192-0093. Specializations: second language acquisition, language teaching methods, teacher education.

Applied Language Learning
1994, Vol. 5, No.1, pp. 73-91

Pragmatics of Interaction: Expressing Thanks in a Second Language[1]

Eli Hinkel
Xavier University

Research into L2 (second language) learning has established that NNSs (Non-Native Speakers) frequently rely on L1 (native language) cultural notions of appropriate behavior in their L2 interactions. Even though this pragmatic transfer can often lead to miscommunication, many commonly performed L2 speech acts are overlooked in language training curricula. A speech act as ubiquitous as expressing thanks is one example. Advanced non-native speakers may perceive the speech act of giving thanks as one of incurring indebtedness or asserting a reciprocal social relationship, as is often the case in their L1s, whereas native speakers of American English perceive it as neither. To investigate whether NNSs with an extensive exposure to L2 culture judge the appropriateness of speech acts of giving thanks in ways similar to those of NSs (Native Speakers), a written role play was administered to 199 highly advanced NNSs. The results show that speakers of Chinese, Indonesian, Korean, Japanese, Spanish, and Arabic ranked expressions of thanks for appropriateness substantially differently than did NSs. While there was a high concurrence about appropriate expressions of thanks within various language groups, there was almost no correspondence among the judgments of different language groups. One implication of this finding is that there may be few aspects of the acquisition of L2 pragmatics that can be excluded from a language training curriculum.

Introduction

Giving thanks is a speech act that is acquired at an early age (Becker & Smenner, 1986) and commonly performed (Coulmas, 1981). For this reason, it is often assumed that learners can successfully say *thank you* in a second language. Methodologies for teaching ESL frequently overlook the substantial cross-cultural differences associated

© 1994, Eli Hinkel

with relatively "simple" aspects of L2 pragmatics, such as expressing thanks. One's ability to say *thank you* in a second language does not evince one's ability to recognize the situations when this speech act is appropriate or the extent to which thanks should be given, according to the rules of politeness in an L2 community.

Most ESL teachers have observed common situations in which speakers of Arabic, for example, may use thanks inappropriately, as in (1) and (2): [2]

(1) A: *I am sorry, I can't talk to you right now. I am on my way to a meeting, but we can talk tomorrow.*
B: *Thank you very much. I'll see you in class. Thank you very much.*

(2) A: *I'll write down the time of the test for you.*
B: *Thank you very much. You are so kind. Thank you.*

In such interactions, many NSs would probably conclude that NNSs did not know how to respond appropriately or perceived A's behavior to be particularly generous. Similarly, in interactions with Indonesians, Koreans, and the Japanese, NSs may think that NNSs didn't thank them because they didn't understand the extent of assistance provided, as in (3) and (4):

(3) A: *I could help you edit your thesis. Maybe I can finish it before the deadline.*
B: *Okay. I really need to meet the deadline.*

(4) A: *If you cannot find an apartment, you could stay at my house for a week or two.*
B. *Great. It'll give me time to keep looking.*

While B's enactment of thanks in (1) - (2) can be viewed as exaggerated and somewhat ingratiating , the lack of thanks in (3) - (4) may be damaging because failing to express thanks when it is expected is unlikely to prompt additional offers of assistance.

Yoon (1991) examined the usage of *thank you* in responses to compliments by monolingual speakers of American English and Korean, and bilingual speakers of both languages. The author's findings show that the bilingual subjects employed politeness strategies divergent from those of monolingual speakers of English or Korean. Eisenstein and Bodman (1986) have similarly determined that L1 sociolinguistic behaviors and L2 grammatical and lexical constraints affect learner production of appropriate pragmatic forms for expressing thanks in a second language.

This study proposes that there may be few aspects of the acquisition of L2 pragmatics that can be excluded from the language-training curriculum. The purpose of this study is to investigate whether advanced and trained non-native speakers, with an extensive exposure to L2 culture, judge the appropriateness of giving thanks in common situations similarly to native speakers of English.

The Cross-Cultural Pragmatics of Giving Thanks

Apte (1974) notes that in many South Asian languages, verbalization of thanks and indebtedness are very closely interrelated, and saying thank you often entails an obligation not only for the speaker, but also for members of the speaker's kin. Similarly, Lebra (1976) indicates that the Japanese may be reluctant to express thanks verbally for fear of incurring indebtedness. In Japanese interactions, the cooperative conversational burden placed on the speaker is considerable as he/she needs to take into account both his/her and the addressee's indebtedness.

Crane (1978) finds that Koreans have a similar cultural precept in which expressing thanks implies taking on obligations. The author also mentions that Koreans tend to be as careful as the Japanese when accepting favors, promises, assurances, or material gifts. On the other hand, Korean expressions of thanks can also carry an element of a reciprocal social relationship, which is not the case in Japanese. Yang (1986) mentions that in Korean culture, social status determines how one expresses thanks. In his example, individuals with a high social position are not expected to give thanks as readily as those whose social status is lower.

As in Korean social frameworks, in Indonesian communities the extent of linguistic politeness one must employ is inversely related to one's social status, power, and control; i.e., the higher one's social standing, the less linguistic politeness the person uses (Errington, 1984; Smith-Hefner, 1988). Because one's perceived social standing is often bound to self-identity, gender, and age, a very complex and elaborate system of tacit rules governs linguistic politeness and, therefore, the giving of thanks (Smith-Hefner, 1988).

In Chinese, expressing thanks occupies a comparatively low place in the hierarchy of politeness, and giving thanks does not entail indebtedness or social reciprocity (Lii-Shih, 1988). Unlike in Korean and Indonesian societies, in Chinese culture, the speaker's and the addressee's age is the primary factor in determining the degree of politeness the speaker employs; e.g., even classmates and friends who are older are treated with respect according to the age difference (Lii-Shih, 1988).

On the other hand, in some cultures, such as Arabic ones, giving thanks is utilized as a means of establishing a closer social relationship: Expressing thanks establishes social reciprocity and group membership

between the speaker and the addressee (El-Sayed, 1990). In fact, Arabic has a special expression to signal when indebtedness is incurred: If the speaker uses the expressions *L'azhli* or *al'asheni*, "for my sake," he or she indicates his or her indebtedness; in most situations, however, indebtedness is not incurred. The etiquette of social interactions requires that offers and promises be made and thanks be given for these offers and promises. The purpose of such exchanges, however, lies in the manifestation of one's being accepted by a social group, rather than in the actuality of the offers and the thanks expressed in response (El-Sayed, 1990; Nydell, 1987).

In Spanish-speaking cultures, norms of appropriateness in linguistic politeness differ strongly between males and females. Several experiments have demonstrated that these differences were so great that no relationship could be shown between male and female perceptions of appropriateness in politeness strategies in Spanish (Becker & Smenner, 1986; Devine, 1982). Devine (1982) found that among various Spanish-speaking groups, e.g., Venezuelan, Mexican, and Puerto Rican, males are expected to say more than is actually necessary in order to be polite. In expressing regrets and thanks, exaggeration is a common politeness reinforcement used by both males and females to ensure sincerity (Brown & Levinson, 1987).

Yet, in other cultures, as in most English-speaking cultures, giving thanks rarely entails indebtedness or establishes social reciprocity. Thanks can be given regardless of whether the offer or promise is accepted or rejected and can be utilized as an acknowledgement (Hymes, 1971). In such non-debt-sensitive cultures, thanks can express gratitude, an intention to express gratitude, or fulfill a social expectation that gratitude be expressed (Hinkel, 1992).

Dialect variations even within English-speaking societies make for substantial differences in how thanks are used or given. Hymes (1971) notes that in British English, giving thanks is different than in American English, with *thank you* frequently serving as a a discourse marker or conversation sequencer, rather than as an expression of gratitude. In neither American nor British English do the rules of politeness governing expressions of thanks generally include such considerations as indebtedness, social status and reciprocity, gender, or age.

621

The Study

Subjects

All participants of the study were enrolled at the Ohio State University and were actively pursuing studies toward their graduate or undergraduate degrees. Of the 233 students, 98 were speakers of Chinese (CH), 33 of Indonesian (IN), 25 of Korean (KR), 15 each of Japanese (JP) and Spanish (SP), and 13 of Arabic (AR). There were 34 NSs included as a control group.

The NNSs represented a highly advanced group of language learners with a mean *Test of English as a Foreign Language (TOEFL)* score of 593. Their residence in the United States typically fell within the range of 1 to 5 years, with a mean of 3.7. It follows that all NNSs had had relatively extensive exposure to L2 culture and appropriate L2 conversational routines.

Materials

Research in L2 learner perceptions of what represents appropriate communicative routines is limited to observations of conversations (Wolfson, 1986; Takahashi & Beebe, 1987) or collecting data through various questionnaire formats. While discourse completion tests and open-ended instruments are frequently prone to problems associated with interpreting subjects' responses (Rose, 1992) and controlling for peripheral linguistic variables (Rintell & Mitchell, 1989), multiple-choice questionnaires, within limitations, have been proven to be a more effective measure of subjects' judgments of appropriateness (Bouton, 1989; Koike, 1989). This study is based upon written role plays (Takahashi & Beebe, 1987) in a multiple-choice format.

The situations for inclusion in the role-play questionnaire were based on Coulmas' (1981) taxonomy. Coulmas states that the verbalization of thanks is related to some action of a "benefactor" or to an outcome of some action that constitutes the object of thanks, material or immaterial. He cautions that this taxonomy is not definitive, and "other criteria are conceivable" (p. 75). He also stipulates that the taxonomy's classes are not necessarily mutually exclusive.

The situations in the present study were originally solicited from 7 NSs and 36 NNSs (Chinese, Indonesian, Korean, Japanese, Spanish, and Arabic). Of the 43 situations collected during the pilot study, 24 were selected to match Coulmas' taxonomy, 3 per class. (See *Appendix* for the questionnaire situations.) The remaining 19 were considered situationally redundant; e.g., 5 suggestions involved an offer of a soft drink and 3 an invitation to a party.

COULMAS' TAXONOMY OF THANKS BY CLASS

1. thanks *ex ante* (for a promise, offer, invitation)
2. thanks *ex post* (for a favor, invitations (afterwards)
3. thanks for material goods (gifts, services)
4. thanks for immaterial goods (wishes, compliments, congratulations, information)
5. thanks for some action initiated by the benefactor
6. thanks for some action resulting from a request / wish / offer by the beneficiary
7. thanks that imply indebtedness [3]
8. thanks that do not imply indebtedness

Procedure

A peer acquaintance was briefly described in the questionnaire:

When you are responding to the questions, please keep in mind the following imaginary student: KC is a student in your department. You have similar interests in your majors. You have talked to KC several times in the department lounge.

Each item in the questionnaire adhered to the same format:

A situation was briefly described; following the situation, three possible responses were presented. One of the choices contained two expressions of thanks.

(In the discussion below, choices containing two expressions of thanks are referred to as *responses with two thanks;* choices containing no expression of thanks are referred to as *responses with zero thanks.*)

For example:

(1) KC asks you if you would like a soft drink. You can say to KC:

(a) *Thanks. I like this drink.*

(b) *Thanks. I like this drink. Thank you very much.*

(c) *I like this drink.*

(2) KC invites you to a party. After the party, you can say to KC:

(a) *I am glad I came to the party.*

(b) *Thanks. I am glad I came to the party.*

(c) *Thanks. I am glad I came to the party. Thank you very much.*

When responding to each questionnaire item, the subjects were asked to rank the multiple-choice responses. The response which, in their opinion, best fit the situation was marked *1*; the second best selection was marked *2*; and the least appropriate, *3*.

The 24 situations presented to the study participants dealt with everyday events, such as dropping one's glove or carrying packages to the post office. As has been mentioned, the questionnaire contained 3 situations for each class in Coulmas' taxonomy. To make the instrument less prone to cultural bias (Wolfson, 1989), situations were solicited from NNSs, and 3 situations per class were included. In the situations and the accompanying possible responses, all references to gender, age, nationality, and native language were avoided; no pronouns except the first person singular and second person singular were used. The vocabulary utilized in the questionnaire design was limited to high frequency words so as not to impede performance.

Results and Discussion

Because the multiple-choice responses in each questionnaire item formed a three-point continuum (i.e., *responses with zero thanks*, *one thanks*, and *two thanks*), the subjects' ranking of the responses in each questionnaire item were interdependent and cumulative. In this ordinal scale, computations were performed for average rankings of *responses with zero thanks* and *two thanks*, excluding the data for *responses with one thanks*.

The subjects' rankings of the *responses with zero thanks* and *two thanks*, as outlined in the questionnaire, were compiled separately for each class in Coulmas' (1981) taxonomy. The rankings for the 3 situations per class for each language group were averaged, as has become common in L2 pragmatics research (Carrell & Konneker, 1981; Blum-Kulka, 1987; Hinkel, 1994; Koike, 1989).

For example, in the situations with thanks *ex ante* (class 1), the average ranking for the selection with *zero thanks* for NSs was 2.36, Indonesians 1.86, Japanese 2.27, Chinese 2.51, Spanish 2.69, Koreans 2.81, and Arabic 2.92. Relative to the other average values for this item, the Indonesian value is clearly the lowest; i.e., more Indonesians than participants in any other group viewed the *response with zero thanks* as appropriate in thanks *ex ante* situations. Based on these values, Indonesian responses were assigned the rank of 1; Japanese, 2; NSs, 3; Chinese, 4; Spanish, 5; Korean, 6; and Arabic, 7. Average values were similarly ranked for the remaining 7 classes.

To determine the consistency of rankings between each of the subject groups, Kendall's Coefficient of Concordance (*W*) was computed for *responses with zero thanks* in classes 1-8; the same procedure was repeated for *responses with two thanks* (see *Table 1*). When the value of *W* equals 1.00, it indicates perfect concordance (or consistency) of rankings between the language groups. As *W* approaches 0, the responses near randomness. X^2 values were used to determine the significance of *W* coefficients (*Table 1*).

The coefficients of concordance for both *zero thanks* and *two thanks* were very high ($W = .981$ [$p < .001$] and $W = .851$ [$p < .001$] respectively) indicating a strong positive correlation and a high degree of consistency in the subjects' judgment of politeness among different native language groups. These values imply that relative to one another, the subjects in L1 groups displayed similar evaluations of the response appropriateness in questionnaire situations. However, concordance in the judgments among the language groups may or may not entail correlations in their evaluations. For example, if NSs consistently indicate that it is appropriate to thank the person who helps carry packages or type papers and the Japanese consistently show that it is not, their judgments may display concordance relative to one another but have little in common. Similarly, if the speakers of Spanish or Arabic systematically judge the *responses with two thanks* appropriate but NSs view them differently, their opinions, however consistent, may not be congruous.

To establish correlations between the judgments of language groups, Kendall's Tau Coefficients (T) were computed for each pair of average rankings; i.e., NSs and Chinese, NSs and Indonesians, NSs and Koreans, etc. (see *Tables 2* and *3*). Although coefficients of correlations between ranks can be computed in various ways, Kendall's T was selected because it permits comparisons of paired ranks extendible to both partial and multiple correlations.

Table 2 presents Rank Correlation Coefficients (T) for subject's *responses with zero thanks*, and *Table 3* shows this correlation for *responses with two thanks*. As with Kendall's *W*, when the value of T equals 1, it indicates a perfect correlation between the rankings of two subject groups. When the value equals -1, it reflects a perfectly inverse relationship between two variables.

TABLE 1
Kendall's Coefficient of Concordance (N = 233)

Rank Orderings of Responses with Zero Thanks

Class	NSs	CH	IN	KR	JP	SP	AR
No.	k = 34	k = 98	k = 33	k = 25	k = 15	k = 15	k = 13
1	3	4	1	6	2	5	7
2	3	5	1	4	2	6	7
3	3	5	1	4	2	6	7
4	4	3	1	5	2	6	7
5	4	3	1	5	2	6	7
6	4	3	1	5	2	6	7
7	4	3	1	5	2	6	7
8	4	3	2	5	1	6	7

$W = .981$ $\quad x^2 = 48.069$ $\quad p < .001$

Rank Orderings of Responses with Two Thanks

Class	NSs	CH	IN	KR	JP	SP	AR
No.	k = 34	k = 98	k = 33	k = 25	k = 15	k = 15	k = 13
1	6	4	7	2	5	3	1
2	6	4	7	2	5	3	1
3	5	3	7	2	6	4	1
4	4.5	4.5	6	3	7	2	1
5	4.5	4.5	7	3	6	2	1
6	3	4	7	2	5	6	1
7	4.5	4.5	7	3	6	2	1
8	4	5	6	3	7	2	1

$W = .851$ $\quad x^2 = 41.699$ $\quad p < .001$

TABLE 2
Rank Correlation Coefficients (T) for Responses with Zero Thanks

	NSs	CH	IN	KR	JP	SP	AR
NSs	1.00	.84	.07	.00	.07	.50	.61
CH		1.00	.39	.07	.11	.46	.57
IN			1.00	.50	-.14	-.21	.54
KR				1.00	.25	-.21	.07
JP					1.00	.39	-.07
SP						1.00	.14
AR							1.00

TABLE 3
Rank Correlation Coefficients (T) for Responses with Two Thanks

	NSs	CH	IN	KR	JP	SP	AR
NSs	1.00	.86	.43	-.04	-.07	.14	.07
CH		1.00	.57	.00	.00	.61	.07
IN			1.00	.43	.43	-.21	.21
KR				1.00	.57	-.00	.50
JP					1.00	-.07	.57
SP						1.00	-.11
AR							1.00

In *Tables 2* and *3*, most of the T coefficients, other than those in the matrix diagonal, are less than the critical value (.64) ($p = .05$) for a two-tailed test. With the exception of the coefficient between the Chinese and NSs (*Tables 2* and *3*), the T values demonstrate that there is no real relationship between the responses of various language groups. Keeping in mind the noticeable LI cultural differences that exist in the rules for appropriate expression of thanks among the subject groups, these findings come as little surprise.

As has been mentioned, in neither Chinese nor English does giving thanks entail indebtedness or social reciprocity. The significant correlations ($T = .84$ [$p < .01$] and $T = .86$ [$p < .01$] in *Tables 2* and *3*, respectively) between the rankings of the Chinese and NSs for *responses with two thanks* may reflect similarities between the NSs' and the Chinese subjects' perceptions of appropriateness associated with giving thanks. The absence of significant correlation between the T coefficients for all groups of subjects can be explained by several processes affecting the acquisition of L2 pragmatic notions and politeness.

The issue of learner competence and performance represents one of the key issues in adult L2 acquisition. Blum-Kulka (1982, 1983) advanced a hypothesis of L2 pragmatic interlanguage in which learners employ speech acts inappropriate in both L1 and L2. Her premise is based on the notion that learners develop a non-language-specific pragmatic competence which they often employ in production of L2 speech acts. In her view, learners draw on their pragmatic competence and knowledge of L1 speech acts, rules of politeness, and grammar in L2 interactions; if equivalent structures do not exist in L2, learners produce inappropriate L2 pragmatic forms. However, because learners' sociopragmatic behaviors may be also adjusted to incorporate L2 norms of appropriateness, NNS speech acts may be incongruent with both L1 and L2 norms of appropriateness (Blum-Kulka, 1983). However, the high and significant values for consistency among group rankings confirm that the subjects in the same native language group made similar judgments of politeness. Blum-Kulka's (1982) pragmatic interlanguage hypothesis does not appear to hold true in the case of expressing thanks in L2. Because no real correlation was established between their L2 behaviors, it is plausible that subjects transferred their L1 pragmatic competence to L2, rather than exhibiting developmental and/or non-language-specific politeness behaviors.

Some studies point to the possibility that exposure to L2 culture and its pragmatic norms is not always synonymous with experience in it and/or the learning of it. Ioup (1989) and Adamson and Regan (1991) provide considerable evidence that Asian immigrants who have resided in the U.S. for more than 10 years frequently fail to acquire the sociopragmatic norms of their English-speaking communities.

The acquisition of L1 rules of politeness associated with thanks begins approximately at the age of 2 when mothers start urging their children to say *thank you* (Bates, 1981). While learners may develop conceptualizations and cognition of L2 pragmatic rules of expressing thanks, their perceptions of the degree of politeness appear to draw on L1 rules of politeness acquired at a very early age.

Conclusions and Implications for Teaching

This study establishes that in written role-plays, 199 highly advanced and trained speakers of Chinese, Indonesian, Korean, Japanese, Spanish, and Arabic ranked expressions of thanks for appropriateness quite differently than did the NS control group. While between-group values for Kendall's Coefficients of Concordance (*W*) for expressions with zero and two thanks were high and significant, the values for T rank correlation coefficients showed no real relationship between the rankings of subjects with different native language backgrounds; i.e., there was a great deal of consistency in the rankings of appropriate expressions of thanks among subject groups. This finding is consistent with that of Scarcella (1979) who determined that although L2 rapport-building strategies can be acquired in natural interactive situations, in general, socio-pragmatic appropriateness has to be taught.

Inasmuch as people frequently assume that their values and views on appropriateness are shared by others (Ruben, 1987), NNSs may not be aware that they can be perceived as unappreciative or ingratiating if they follow L1 socio-pragmatic norms. Furthermore, because the act of expressing thanks is performed in most cultures, NNSs may not even recognize the differences that exist between their acts of thanks and those of NSs. Learning culturally-appropriate speech act paradigms begins with self-awareness and non-judgmental observation (Gaston, 1984). Teachers can heighten students' awareness of L2 politeness strategies and help them understand the effect of their L1-based behaviors in L2 environments.

Because few culture-training materials exist, both NNSs and their instructors are often faced with an unexplored domain of language teaching, i.e., specific and pragmatically-appropriate formulae associated with many types of speech act. One of the many productive approaches to teaching L2 norms of appropriateness is to encourage students to observe specifically how NSs express thanks (or perform other speech acts), under what circumstances, and to what extent. Their observations can be compared to NNSs' views on appropriateness pertaining to a particular speech act, and the differences and similarities can be discussed and analyzed.

In many ESL programs, the teaching of L2 linguistic politeness norms is rarely addressed in depth. When L2 socio-pragmatics is taught,

it usually becomes a part of the curriculum at advanced levels of student proficiency because ESL methodologies often assume that students at lower proficiency levels may not be ready for the complexity of L2 appropriateness norms. Although expressing thanks is pragmatically complex, linguistically it is not. For this reason, the teaching of appropriate paradigms for expressing thanks can begin at lower levels of student linguistic proficiency and thus introduce students to the importance of developing L2 socio-pragmatic skills.

Notes

1. An earlier draft of this paper was presented at the Sixth International Conference on Pragmatics and Language at the University of Illinois at Urbana-Champaign, 1992.
2. Examples given in this paper come from a collection of observed and recorded data, the conceptual basis for this study.
3. While both NSs and NNSs were requested to submit situations to match Coulmas' (1981) taxonomy, NSs had difficulty with those in which giving thanks implied indebtedness. In fact, several NSs commented that the actual giving of thanks did not entail indebtedness, but rather, the benefactor's actions did.

Appendix

The Questionnaire

1. KC invites you to a party. You accept. You can say to KC:
   (a) *I'll be glad to come to the party.*
   (b) *Thanks. I'll be glad to come to the party. Thank you very much.*
   (c) *Thanks. I'll be glad to come to the party.*

2. You left your calculator at home, and you need to finish your assignment. You ask KC to let you use the one KC has. KC gives you the calculator. When you return it, you can say to KC:
   (a) *Thanks. I finished my assignment.*
   (b) *Thanks. I finished my assignment. Thank you very much.*
   (c) *I finished my assignment.*

3. You are planning on taking a course which KC has already taken. You ask KC about the course. KC tells you about it. You can say to KC:
   (a) *Thanks. I am thinking about taking this course.*

(b) *I am thinking about taking this course.*
(c) *Thanks. I am thinking about taking this course. Thank you very much.*

4. Usually, you walk home. Today, the rain is pouring down and you do not have an umbrella. KC drives home in the direction of your house, and today because of the rain, KC is driving you home. You can say to KC:

(a) I *am glad I didn't have to walk home today.*
(b) *Thanks. I am glad I didn't have to walk home today. Thank you very much.*
(c) *Thanks. The rain is really coming down.*

5. You and KC are walking to class. You drop your glove and do not notice it. KC picks it up and gives it to you. You can say to KC:

(a) *Thanks. I did not even notice I dropped that glove.*
(b) *Thanks. I did not even notice I dropped that glove. Thank you very much.*
(c) *I did not even notice I dropped that glove.*

6. You are going home for the break between quarters. KC wishes you a safe trip and a nice break. You can say to KC:

(a) *Thanks. I am looking forward to going home. Thank you very much.*
(b) *I am looking forward to going home.*
(c) *Thanks. I am looking forward to going home.*

7. You are carrying several large packages to the post office. You are struggling and walking very slowly. KC helps you carry them. When you come to the post office you can say to KC:

(a) *Thanks. I can manage from here. Thank you very much.*
(b) *Thanks. I can manage from here.*
(c) *I can manage from here.*

8. You are typing a paper in the computer lab and having a hard time getting your printer to work. KC gets it to work. You can say to KC:

(a) *Thanks. I can print my paper now.*
(b) *I can print my paper now.*
(c) *Thanks. I can print my paper now. Thank you very much.*

9. You need to write down some information. You do not have a pen. KC hands you one. After you finish writing you can say to KC:

(a) *A pen is just what I needed.*

(b) *Thanks. A pen is just what I needed. Thank you very much.*
(c) *Thanks. A pen is just what I needed.*

10. **You have just received a letter of acceptance to Harvard University. KC congratulates you. You can say to KC:**

    (a) *Thanks. I am very happy. Thank you very much.*
    (b) *I am very happy.*
    (c) *Thanks. I am very happy.*

11. **You are very busy. It's lunch time You are hungry but you do not have time to get food. KC offers to bring you a sandwich from the cafeteria. You accept. You can say to KC:**

    (a) *I am starving.*
    (b) *Thanks. I am starving.*
    (c) *Thanks. I am starving. Thank you very much.*

12. KC asks if you would like a soft drink. You accept. You can say to KC:

    (a) *Thanks. I like this drink.*
    (b) *Thanks. I like this drink. Thank you very much.*
    (c) *I like this drink.*

13. **You are moving to a new apartment, and KC helps carry some things. You can say to KC:**

    (a) *Thanks. I couldn't move this alone. Thank you very much.*
    (b) *I couldn't move this alone.*
    (c) *Thanks. I couldn't move this alone.*

14. **You are in the university bookstore trying to find some books which you need for your course. KC shows you where the books are located. You can say to KC:**

    (a) *Thanks. I didn't know where to look.*
    (b) *Thanks. I didn't know where to look. Thank you very much.*
    (c) *I didn't know where to look.*

15. **You need to go to the other end of the city to take a test (TOEFL, SAT, or GRE). KC knows the city very well. You ask KC for the road directions, and KC explains the route to you. You can say to KC:**

    (a) *Thanks. I've never been to that part of the city. Thank you very much.*
    (b) *Thanks. I've never been to that part of the city.*

(c) *I've never been to that part of the city.*

16. You dropped your papers and notes, which spilled in the middle of a busy sidewalk. KC is walking by. KC stops and helps you pick up your papers. You can say to KC:

    (a) *Thanks. I can't believe I dropped my papers.*
    (b) *I can't believe I dropped my papers.*
    (c) *Thanks. I can't believe I dropped my papers. Thank you very much.*

17. KC gives you a ride that takes about 15 minutes. When you arrive, you can say to KC:

    (a) *It didn't take very long.*
    (b) *Thanks. It didn't take very long. Thank you very much.*
    (c) *Thanks. It didn't take very long.*

18. You ask KC to help you fill out the health insurance form. KC is glad to do it. You can say to KC:

    (a) *Thanks. These forms are hard to fill out.*
    (b) *These forms are hard to fill out.*
    (c) *Thanks. These forms are hard to fill out. Thank you very much.*

19. KC invites you to a party. After the party, you can say to KC:

    (a) *I am glad I came to the party.*
    (b) *Thanks. I am glad I came to the party.*
    (c) *Thanks. I am glad I came to the party. Thank you very much.*

20. Your paper is due in 30 minutes. You still have to finish typing it, and it looks like you may not be done on time. KC helps you with the typing. You can say to KC:

    (a) *It's almost done.*
    (b) *Thanks. It's almost done.*
    (c) *Thanks. It's almost done. Thank you very much.*

21. KC offers to go with you to the department office to help you with some of the procedures for registration. You can say to KC:

    (a) *Thanks. These registration procedures are confusing.*
    (b) *These registration procedures are confusing.*
    (c) *Thanks. These registration procedures are confusing. Thank you very much.*

22. You need to use a book which KC has. KC offers to lend it to you for a few days. You can say to KC:

    (a) *Thanks. I need it. Thank you very much.*
    (b) *I need it.*
    (c) *Thanks. need it.*

23. KC asks you if you would like an apple. You accept. You can say to KC:

    (a) *Thanks. I like apples. Thank you very much.*
    (b) *Thanks. I like apples.*
    (c) *I like apples.*

24. You left your wallet at home and have no money for lunch. KC loans you $5. You can say to KC:

    (a) *Thanks. I'll pay you back tomorrow.*
    (b) *Thanks. I'll pay you back tomorrow. Thank you very much.*
    (c) *I'll pay you back tomorrow.*

References

Adamson, H. & Regan, V. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language: A preliminary study. *Studies in Second Language Acquisition, 13*, 1-22.

Apte, M. (1974). "Thank you" and South Asian languages: A comparative sociolinguistic study. *International Journal of the Sociology of Language, 3*, 67-89.

Bates, E. (1981). *Language and context: The acquisition of pragmatics.* New York: Academic Press.

Becker, J. & Smenner, P. (1986). The spontaneous use of *thank you* by preschoolers as a function of sex, socioeconomic status, and listener status. *Language in Society, 15(4)*, 537-546.

Blum-Kulka, S. (1982), Learning to say what you mean in a second language: A study of speech act performance of Hebrew second-language learners. *Applied Linguistics, 3*, 29-59.

Blum-Kulka, S. (1983). Interpreting and performing speech acts in second language: A cross-cultural study of Hebrew and English. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and second language acquisition.* Rowley, MA: Newbury House.

Blum-Kulka, S. (1987). Indirectness and politeness in requests: Same or different? *Journal of Pragmatics, 11*, 131-146.

Blum-Kulka, S. (1989). Playing it safe: The role of conventionality in indirect requests. In S. Blum-Kulka, J. House, & G. Kasper

(Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 155-173). Norwood, NJ: Ablex.

Bouton, L. (1989). So they got the message, but how did they get it? *IDEAL, 4*, 119-148.

Brown, P. & Levinson, S. (1987). *Politeness*. (Reprinted) Cambridge: Cambridge University Press.

Carrell, P. & Konneker, B. (1981). Politeness: Comparing native and non-native judgments. *Language Learning, 31*, 17-30.

Coulmas, F. (1981). Poison to your soul. Thanks and apologies contrastively viewed. In F. Coulmas (Ed.), *Conversational routine: Explorations in standardized communication situations and pre-patterned speech*. The Hague, Netherlands: Mouton.

Crane, P. (1978). *Korean patterns*. Seoul, Korea: Kwangjin Publishing.

Devine, J. (1982). A question of universality: Conversational principles and implications. In M. Clarke & J. Handscombe (Eds.), *On TESOL '82: Pacific perspectives on language learning and teaching*. Washington, DC: TESOL.

Eisenstein, M. & Bodman, J. (1986). "I very appreciate": Expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics, 7*, 157-185.

El-Sayed, A. (1990). Politeness formulas in English and Arabic: A contrastive study. *ITL, 89-90*, 1-23.

Errington, J. (1984). Self and self-conduct among the Javanese priyayi elite. *American Ethnologist, 11(2)*, 275-290.

Gaston, J. (1984). *Cultural awareness teaching techniques*. Brattleboro, VT: Prolingua.

Hinkel, E. (1992). *Appropriateness in L2 conversations: Function and form*. Paper presented at NAFSA 44th Annual Conference, Chicago, IL.

Hinkel, E. (1994). Topics appropriate in cross-cultural social conversations. *Pragmatics and Language Learning, 5*, 163-179.

Hymes, D. (1971). Sociolinguistics and the ethnography of speaking. In E. Ardner (Ed.), *Social anthropology and language*. London: Tavistock Publications.

Ioup, G. (1989). Immigrant children who have failed to acquire native English. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition*, Vol. 2. Philadelphia: Multilingual Matters.

Koike, D. (1989). Pragmatic competence and adult L2 acquisition: Speech acts in interlanguage. *Modern Language Journal, 73(3)*, 279-289.

Lebra, T. (1976). *Japanese behavior patterns*. Honolulu: University Press of Hawaii.

Lii-Shih, Y. (1988). *Conversational politeness and foreign language teaching*. Taipei, Taiwan: Crane Publishing.

Nydell, M. (1987). *Understanding Arabs.* Yarmouth, ME: Intercultural Press.

Scarcella, R. (1979). On speaking politely in a second language. In C. Yorio, K. Perkins, & J. Schachter (Eds.), *On TESOL '79: The learner in focus* (pp. 275-287). Washington, DC: TESOL.

Rintell, E. & Mitchell, C. (1989). Studying requests and apologies: An inquiry into method. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 155-173). Norwood, NJ: Ablex Publishing.

Rose, K. (1992). Speech acts and questionnaires: The effect of hearer response. *Journal of Pragmatics, 17,* 49-62.

Ruben, B. (1987). Guidelines for cross-cultural communication effectiveness. In L. Luce & E. Smith (Eds.), *Toward internationalism* (pp. 36-48). Cambridge, MA: Newbury House.

Smith-Hefner, N. (1988). Women and politeness: The Javanese example. *Language in Society, 17(4),* 535-554.

Takahashi, T., & Beebe, L. (1987). The development of pragmatic competence by Japanese learners of English. *JALT, 8,* 131-155.

Wolfson, N. (1986). Research methodology and the question of validity. *TESOL Quarterly, 20,* 689-699.

Wolfson, N. (1989). Problems in the comparison of speech acts across cultures. In S. Blum-Kulka, J. House, & G. Kasper (Eds.).

Yang, W. (1986). *Korean ways, Korean mind.* Seoul: Tamgu Dang Publishing.

Yoon, K. (1991). Bilingual pragmatic transfer to speech acts: Bidirectional responses to a compliment. In L. Bouton & Y. Kachru (Eds.), *Pragmatics and Language Learning, Vol. 2* (pp. *75-100).* Urbana-Champaign: University of Illinois.

Author

ELI HINKEL, Director, ESL Program, Xavier University, Cincinnati, OH 45207. Specializations: pragmatics, second language learning, curriculum design.

636

Applied Language Learning
1994, Vol. 5, No.1, pp. 93-114

Deviant English and the Para-Poetic

Clark Lunberry
Suzugamine Women's College

in memory of John Cage

Foreword

Most journal articles which examine language-learner discourse scientifically analyze data or gradual change as the learner pursues native-like behavior. Clark Lunberry's article, while describing "errors" in traditional terms on one hand, encourages unusual and interesting views of error-filled writing on the other. His description of learner writing as "para-poetic" suggests that the instructor-reader look beyond conventional tense and aspect, agreement and article to see the imagery of unusual juxtaposition of words and phrases in student work. Whether or not the students intend to be poetic, the teacher could certainly find his work more enjoyable in taking Lunberry's advice to see what is there rather than criticizing what is not.

Introduction

Uttering a word is like striking a note on the keyboard of the imagination.

— Ludwig Wittgenstein

You don't yourself until now from tomorrow.

Forming a part of a short essay describing a dream, the above sentence is an example of the kind of writing frequently encountered in the teaching of English as a foreign language. Non-native students of English will often speak and write in ways that, in addition to being technically incorrect, seem both peculiar and provocative to the native speaker. The individual words are often recognizable, but the awkward manner in which they have been combined creates obstacles that render their meaning anywhere from interesting to incomprehensible. The native speaker detects linguistic disruptions almost immediately,

© 1994, Clark Lunberry

possessing a seemingly instinctive awareness of conventional boundaries and the limits of language. Like a tongue exploring the roof of its own mouth, abnormalities are quickly felt and focused upon. However, regardless of the errors, regardless of the awkwardness, there exists a fleeting moment in the mind of the native speaker, caught off-guard, when the words and sentences register and resonate with potential meaning. The words join together and form an obscure, opaque pattern. The sentence, *You don't yourself until now from tomorrow*, though vague and unsettled, incites fragile possibilities and strange suggestion.

Like Wittgenstein's keyboard of the imagination,[1] language possesses rich musical qualities of tone and texture, resonance and rhythm. These more mysterious, less quantifiable elements of language are key factors in determining everything from nuance to nonsense. However, textbooks and dictionaries, language classes and language teachers generally present only the most contrived and standardized forms of English. They ignore, perhaps inescapably, the more subtle and intangible qualities of cultural context and linguistic diversity.[2] In the process, what the students receive is a decontextualized, abstract language which, in their eyes, is primarily an intricate and obscure web of words and phrases, rules and regulations. What they fail to hear is the music of language, the visceral quality of words to strike and resonate within the imagination.

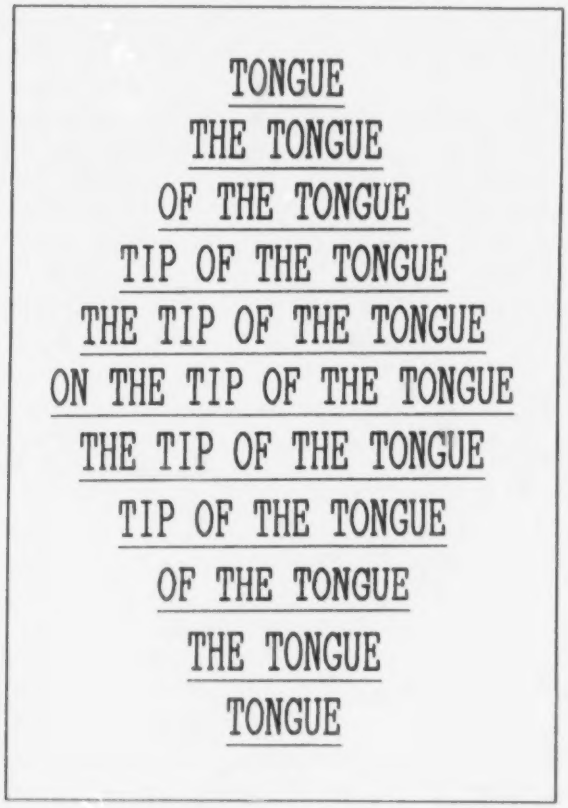
Imagine trying to teach students to swim by using textbooks, diagrams, explanations on the blackboard, and various other classroom methods. The students may well understand all that has been said, scoring high on essays and examinations. This manner of teaching might become so institutionalized and accepted that the very real action of swimming in water is nearly forgotten. But then the day comes for the students to be thrown into the pool—the real test! Will they sink or swim?

The non-native students of English may find themselves in a similar, though perhaps less dramatic or dangerous situation.[3] They are suddenly thrown into the cold water of English. For the lower-level student, various maneuvers are made—a preposition thrown here, a modifier added there, articles and adjectives tossed about everywhere. When written out as an essay or examination, the choice of words may seem random, the use of grammar all but arbitrary, awareness of tone and rhythm nearly nonexistent. And yet, mysteriously, paradoxically, they remain afloat. Their words and sentences, however mangled and confused, still manage somehow to convey traces of meaning and fragments of image.

The native speaker reads the writing of such students with a combination of horror and fascination. Perched on the very edge of nonsense and incomprehensibility, the student has alarmingly pulled off

638

a remarkable feat of language. Unknown to its own creator, the deviant English is possessed of a power that shakes the foundation of the native speaker's carefully constructed house of language. What their writing inadvertently demonstrates is that in spite of the formalized and contrived English that they have been taught, in spite of the errors and irregularities that they employ, the language itself cannot be extinguished and rendered meaningless. The inherent power of individual words and fragments to *mean*—to strike a note on the keyboard of the imagination—reveals the limits and fallibility of linguistic rules and testifies to the latent power and vitality of an expanded, outlaw language.



639

Clark Lunberry

The Para-Poetic Inter-Viewed

Somehow it seems
to fill my head
with ideas –
though I don't
exactly know
what they are.

— Alice in Wonderland

Q: How would you define the para-poetic?[4]

A. The para-poetic is any use of language that accidentally or unconsciously employs words or sentences in interesting and unexpected arrangements. With the para-poetic, rules of correct language use are often violated, resulting in unusual formations of grammar, word-choice, syntax, punctuation, spelling, etc. However (and significantly), these violations are not deliberate transgressions but instead an unintentional straying beyond boundaries otherwise unknown.

Q: Who writes para-poetry?

A: The most interesting examples come from those who fall between languages. For my purposes I have focused upon the writing of non-native, low-to-intermediate level Japanese students of English. Their weaknesses with the language, their lack of poetic pretension, combined with the radically different structure of the Japanese language, often generate extraordinary inter-lingual effects.

Q: How is para-poetry written?

A: Varying according to individual ability, the students' level of English generally allows a limited and insecure grasp over what they are trying to say. Their bond with the mother tongue remaining strong, they commonly attempt to transfer thoughts directly from one language into another. Though discouraged from doing so, the students often pre-write their essays in their native language. Later they go back through with a Japanese-English dictionary in the frenzied but feeble hope of finding approximate equivalents. Not surprisingly, in this situation and for their purposes the dictionary does more harm than good. It is clearly not the bridge between languages that they had intended it to be.

Q: Do the students know that they have been para-poetic?

A: No! Understanding some of the particulars, they miss the overall tone and texture. As suggested before, they are generally not

640

attempting to be poetic or particularly creative. They are in most cases simply struggling to use the English they possess while fulfilling a modest homework assignment. It is precisely through the absence of creative intention that they succeed. They are not using language poetically, but rather language is using them para-poetically! They are not innovators, they are not poets; they are vectors of influence through which language is filtered and scattered about—surrogates whose brains and bodies are used to conceive otherwise unforeseen results.

Q: *What makes the writing "para-poetic" and not simply bad English?*

A: Interestingly, it is both. In fact, it is precisely through its badness that it is good. As it fails on one level, it succeeds on another—presenting forms of language that, though technically incorrect or conventionally awkward, remain strange and evocative. For the native speaker who reads a student's writing, the reassuringly familiar nature of language is quite radically disrupted. Yes, it is full of errors, an affront to "fine writing," but through its benign disregard of linguistic and poetic conventions a remarkable freshness and aberrant vitality emerges.

Q: *Is it really possible to write poetically without trying to—in other words, to "be a poet and not know it?"*

A: Yes! In fact, the opposite situation is often a much greater obstacle—the poet trying too hard, full of intention, sound, and fury, signifying too much and nothing simultaneously. The weaknesses and whims of individual ego so often produce, at best, a self-indulgent boast or cloying confession. Language is arrogantly assumed to be a tool that one controls towards some kind of precise description of a delicate inner event. Perhaps it is just as much the contrary that is true: We are the tool and language is in control of us. And besides, language probably has more to say than we ever could. We get in the way of allowing language to speak![5] The para-poet doesn't for a moment forget the immensity and strangeness of language. Never boasting any mastery, the non-native student remains humble and beholden before the power and complexity of this impenetrable network of signs. The student does not presume to manipulate its meanings and is instead, either consciously or unconsciously, receptive to language's manipulation.

Q: *What is the relation of the para-poetic to the more conventional language of the native speaker?*

A: Language is a historical repository of inherited forms of communication and expression. Born of context and situation, language emerges from the needs and desires of a culture and a

community. It is larger than the individual because it precedes the individual. Absorbed and assimilated from birth, *we* are precisely our own linguistic limits and language far exceeds our own capacity to use it. We breathe our language as a vital, sustaining element—a life-giving force. Language is a thick, pressurized substance, an atmospheric event comprised of words, expressions, metaphor, rhythm, tones, tropes, textures, tenses, noises, nuance, and nonsense. Inevitably and necessarily, the culture that produces language also confines it. Left to its own devices, it can quickly break through boundaries meant to hold it in. Inherently promiscuous, language is divided by a fragile barrier separating carefully contained expression from the perverse screams of a madman. Belonging to another language and another culture, it is this delicate boundary that the non-native student is not in a position to fear or respect. They cross the line of correct, conventional language use not out of courage or curiosity, but out of ignorance, like a child innocently straying beyond the fenced-in playground out onto the raging freeway.

NANCY reprinted by permission of UFS, Inc.

98
BEST COPY AVAILABLE

Where Are We? When Are We?

Language is a labyrinth of paths.
You approach from one side
and know your way about;
you approach the same place
from another side and
no longer know where you are

—Wittgenstein

Before presenting a collection of student writings I would like to make a more detailed examination of two individual examples of the para-poetic—one full of errors, the other containing only a few. Perhaps a more concrete and specific analysis will yield some understanding of what makes the para-poetic, poetic.

The following piece of writing was done by the 19-year-old Japanese college student Shiho Takahashi. The assignment was to write a paragraph describing sounds heard at night.

> *One night when my parents are going on a trip, I'll go to bed in my room. Then I heard sound of window's broken from the first floor. I was very surprised and went down stairs with fearful. Then clock has just struck twelve. In outside, a dog is howling and I heard siren of a police car. I got weirder and weirder. I opened door of living room. Then the phone has just rung.*

This paragraph stands as a striking example of the para-poetic. Along with the many incorrect and incongruous elements contained within it, perhaps the most striking and disorienting feature is the student's use—or abuse—of verb tense. Taken so much for granted by the native speaker, it remains a thorn in the side of the non-native student even after years of study. However, to incorrectly or inconsistently conjugate the verbs describing an event is to deny a solid foundation from which that event can be imagined. Verb tense has the fundamental and, one could say, metaphysical function of situating the subject in time, structurally sustaining frames of temporal perception. Lacking a strong and unshifting point of view, any and all action fails to settle and cohere upon an orderly time line. This violation of linear development presents a radically unstable stage upon which seemingly little intelligible or coherent could occur. And yet something does occur. Regardless of the errors and irregularities, language is not neutralized or rendered impotent. Though employed with a degree of blindness, language still functions as a generator of images and meanings. With

the para-poetic, events don't simply unfold, they unravel in ways that defy and challenge conventional methods of description.

With a focus on verb tense, look again at Shiho Takahashi's piece of writing. Suspend for a moment your demand for linguistic correctness. Forget this was written by a student. Forget that it was written in response to a simple homework assignment. Treat this piece of writing as an example of language existing beyond "good" or "evil." Follow her through the description just as it is provided, move with each of the twisting tenses, allow yourself to be swept about in the chaotic flow of words and images.

Shiho Takahashi begins with the temporal designation "one night." From this, the reader immediately locates the action, assuming that the events about to unfold will occur in the past. However, midway through the sentence we are suddenly shifted into the present continuous as we hear that the parents are "going on a trip." Already wavering between two times (Where are we? When are we?) she then proposes a third—announcing in the future tense that she "will go to bed." Within one sentence, we find ourselves inhabiting three distinct time zones. As if that isn't confusing enough, she then begins her next sentence by disorienting us even further—dropping us solidly back into the simple past with "then I heard." She proceeds along in this manner through two sentences and we quickly adjust to our placement in the past, perhaps thinking that we've finally settled. But no! With the next sentence we are pulled back into the present with the present perfect (appropriately enough with the striking of a clock). Settling for a moment in the present where a "dog is howling," we then slide back into the past with the "heard siren of a police car." The action proceeds in the past and we seem again to be settling but only to be struck by the final *coup de grace*: a sudden acceleration towards a conclusion in the present perfect—"the phone has just rung."

The events of this story unravel as though seen on a video screen. However, instead of a linear development of action, one watches the story with a remote control gone haywire. We push the *play* button and the action begins. As the story progresses, the *rewind* button is accidentally hit. In a panic to readjust, the *fast forward* is pressed. However it's held too long so when *play* is reactivated we've gone too far. Out of frustration we hit *pause*, and stop all the action . . . take a breath, then once more, the *rewind*. Overshooting again, we hit *play* and for a moment rewatch a part of the story already seen. Recognizing it as the past, we again fast-forward and this whole process goes on and on and on—a desperate search to relocate the vanished but remembered present, the point in the story that has been so frustratingly lost.

In addition to the subject of tense and the subsequent confusion of temporal location, an analysis of this writing also demonstrates an equally disorienting, and equally interesting, depiction of space. The

most obvious and intriguing example is the student's use of the technically incorrect combination "in outside" from the sentence "In outside, a dog is howling." Though "incorrect," the use of the preposition *in* before *outside* is a clear and fascinating example of the para-poetic. Presumably derived directly from the Japanese equivalent *soto de*, the student has perhaps speculated that because in Japanese the word for *outside* (soto) is grammatically joined by the preposition for *in* (de), a literal translation should reproduce both words. However, with English, placing *in* before *outside* strangely and evocatively acts to mentally confine and place boundaries around a conceptual space that in standard linguistic terms must be left unbounded. For the native English speaker, the word *outside* immediately conjures up a boundless spatial range.[6] When one is "outside," one is necessarily not "in" anything. By modifying *outside*, a conflict immediately arises as a finite and more specific space is envisioned. With conventional language use, being "in outside" is simply not permitted and the conscientious English teacher considers it incorrect. The expression is withdrawn from circulation as though violating some immutable laws of linguistic logic, or even more, challenging undefined metaphysical concepts of space. However, as para-poetry, *in outside* is permitted and, as anyone with an open ear to language will know, this simple, two-word, "incorrect" combination succeeds in evoking a conceptual image otherwise denied by conventional, correct language.

FIRST LANGUAGE = MOTHER TONGUE

SECOND LANGUAGE = ?

Definitely Indefinite / Indefinitely Definite

Words strain,
Crack and sometimes break under the burden,
Under the tension, slip, slide, perish,
Decay with imprecision, will not stay in place,
Will not stay still.

— T. S. Eliot

645

The second example of student writing is of a very different nature from the previous one. It was written by a 19-year-old Japanese college student, Kaori Miura, whose technical ability is remarkably high: Her imaginative description of a dream is impressive. As you will see, she has little problem with verb tense, remaining consistently anchored in the past. In fact, this writing just barely passes as para-poetic. It is almost too good—there are very few errors and the power of the student's images is seemingly not accidental. However, the errors it does possess are forceful, transforming the intended meaning of essential details. Demonstrating a more subtle manifestation of the para-poetic, an already "good" essay unconsciously engenders a deviant variation of itself.

I was studying in room. I was using dictionary
to look up word meaning. Suddenly, word disappeared.
I was surprised and I turned over two or three pages.
All page's word disappeared. Then, red spot appeared in
the white pages. It became bigger and bigger. At that time
I woke up. And I searched for my dictionary. I found
the last page was stained with red spot. And I took the
other book and turned over the pages.
All page's word disappeared.
Then I looked at myself in the mirror.
All the word was sticked in my face
I shouted and I went down the stairs
in a hurry. I went to my parent's bedroom. I shouted
again. My parent's face was pitch-black.

Kaori Miura's description of a dream is vivid and powerful. Like so many dreams, almost everything is precisely as it should be—familiar people, familiar objects, familiar rooms. What stands out for the dreamer—what stands out for the reader—are the incongruous details, minor alterations that disrupt habitual expectations. In this piece of writing, the para-poetic element is seen in the subtle descriptive features, small errors that create vagueness and cause confusion. The inherent strangeness of a dream is subtly enhanced through accidental slips of linguistic precision.

One of the recurring and interesting errors in this essay involves the use, or misuse, of definite and indefinite articles. Like verb conjugation, the application of articles before nouns is very much second nature to a native speaker of English, but remains a problem for the non-native student. This is especially difficult for many Japanese students because their language essentially lacks anything comparable to an article. They are therefore obliged to develop and deploy a linguistic form for a concept that their language does not possess. Further, in many ways the

Japanese language exhibits a determined and desired imprecision that conflicts with the detailed and precise descriptions afforded by articles, reflecting a cultural attitude as much as a linguistic one. There is a kind of built-in resistance to the more rigid, tightly defined frame of reference imposed by English. In fact, students are often both puzzled by articles and reluctant to bother with them at all; a student once told me that "it's not important" whether articles, seemingly so insignificant, are used or not. She continued, "You understand what I mean even if I forget *a* or *the*."[7] Of course, this student is partially correct. We do basically understand what is meant when we read something like, *I was using dictionary to look up word meaning*. What the protesting student fails to realize is that the absence of an article is immediately, almost physically, felt by the native speaker—like having a small but cherished piece of furniture suddenly vanish from a familiar room.

In addition to this rhythmic irregularity, an even more subtle disruption of meaning is produced. For example, to say "I was using dictionary to look up word" is to leave vague a noun's precise position and function. The purpose of articles, both definite and indefinite, is to signal a noun and specify its intended application. Failing to use them correctly creates an ambiguity in which the status of a noun remains obscure. To say *word* but not *a word* or *the word* is to evoke an odd, generalized image possessing for the native speaker subtly different connotative potential. Using a dictionary to "look up word" could suggest a possible search for all words, words-in-general, an abstract condition of *word*-ness that transcends the particular and evokes the idealized. Granted, such bizarre implications are generally not seriously, consciously considered by the native speaker, but on an unconscious level they strike a note that inevitably generates fragments of image and moments of meaning. If only for a fraction of a second, error is not read as error but is instead registered in the mind as possibility, received, imagined, and then quickly—too quickly—dismissed as technically incorrect and therefore undeserving of consequence or attention. The logic of linguistic organization reaffirms itself and the aberrant English soon dissipates into nonsense. However, it is from this darker realm of language's excess,[8] in that fleeting, imaginative moment prior to the restoration of order, that errors and inconsistencies emerge and the para-poetic is born.

Further reinforcing the general, unspecific nature of the noun is the student's use of the singular for what is presumably plural. With the sentence *All page's word disappeared*, the more conventional, appropriate manner of writing would perhaps be *All the words on the page disappeared*. The conscientious language teacher would correct the sentence by rearranging the word order, inserting the definite article, making plural the singular, more or less insisting upon this one form of writing.

An additional rich example of singular/plural confusion can be seen in the final sentence of the paragraph, *My parent's face was pitch-black*. By failing to make plural her parents' faces, the student creates a compelling image in which two logically distinct elements have been conceptually fused into one. Manifesting a kind of Freudian condensation, the student convincingly represents the unstable mental processes and disrupted perception of an actual dream.[9] Within a dream, it is entirely plausible that the faces of parents would in fact be condensed and experienced as one. The student's "error" represents a fusing of the linguistic and the psychic, inadvertently but convincingly depicting a dream's visual presence.

However, the precise and inflexible logic of conventional language would once again dictate a more *technically* accurate description, providing the parents with *faces*, not *face*. The student would be required to correct the situation, dispel the vagueness, and return us to the rational precision of a stable, distinct depiction. But in the process, the rich and compelling resonance of the parent's singular *face* would be lost, the linguistic manifestation of dream condensation denied.

The Poetless Poem

One great part of every human existence is passed
in a state that cannot be rendered sensible
by the use of wideawake language,
cutanddry grammar and
goahead plot.

— James Joyce

In conclusion I will now present a much larger group of para-poetic writings. The writings were collected from assignments by Japanese college students and have been divided into six groups, determined loosely by subject matter. Each of the six sections has been given a title taken from a para-poetic fragment within the group. In the two examples of student writing already discussed, numerous factors have been pointed out as contributing to form the para-poetic. With the first example there was the inconsistent use of verb tense and the unconventional application of a preposition. With the second example it was a confusion of the singular and the plural and the irregular use of definite and indefinite articles. These are only a few instances of the kinds of errors frequently encountered with student writings and in the pages that follow numerous others will be seen. However, the point of para-poetry is not "error analysis" but rather a kind of error enjoyment; a pleasure derived from a playful, liberating, inadvertently poetic application of language. Presented from here on out without comment,

648

the students' writings have been reproduced as faithfully as possible, maintaining the errors and irregularities as they were originally found.

Sounds Me Hear

Sounds me hear from a window is many sounds every days.
Sounds of raining, it makes me melancoly lonely
and become wanted to see someone.
Sounds of wind, it gives me fine and become a fresh
and become wanted to go out.
Sounds of songs of birds I get up as if alarm clock.
My feelings is simple as if be influenced by sound.

— Yoko Higashi

If I were not hear a sound,
I can't quite spread my family's voice, my friend's voice,
town's voice, bird's voice and some tree's voice.
Words are not adequate to express my feeling.
I don't know I should do.
I can't hear a sound now, till I have been heard
the music from radio, and I'll quite hear my favorite music.
It is very stuffy, when I think this problem.
But deaf's man is living such a life.
They are cheerful, and they are like life with the
general public.
I must be able to approach with their, talk with the
sign and communication.
So We are healthy and normal.
We should communicate us thinking for their.

— Ryoko Asano

If I were deaf I would missed different other people.
Everyone can hear everything but I can't hear it.
I'm afraid if I were such a think.
If I were deaf I would not spoken anything.
listened to music, heared the voice of my parents
and my friends and understood what happen to now.
I may mad. I may riot.
But it may not nothing but bad.
I will feel kind and help my around people.
I want to meet people who understand my mind.
—people who think all right if I don't speak—

Also I want to effort to speak.
I think we can understand each other more than now.
because we express by all body.
It is hard for me to live the rest of life
but I will find importance to live.

— Noriko Nakagami

I hear a variety of sound. The singing of a bird.
A crow caws. The sound of alarm. The sound of down
crossing gate. The sound of go train. The sound of
run car. The sound of run motor bicycle. The chatting
of neighbor. The sound of sound a horn. The sound
of television from next room. The sound of do the
cooking. The sound of flying jet plane. The chirping
of an insect. The sound of pedal bicycle.
The chat, cry, laugh, sing, play a school child.
The sound of pump, sneaker and so on.
The bark of dog. The sound of ringing the telephone.
The sound of ambulance. The sound of radio.
The sound of open the door. The sound of sweeping yard.
The sound of foliage when the wind is blow.
The tolling of a bell from temple.
It has been five o'clock.

— Kaori Takeda

Strange Person Everyone

I'm often spoken to a strange person. I think that why?
But I'm a strange person that the person see to me.
So, the people all over the world are strange person everyone.
But I think that a strange person is odd.
A strange person is not a laughable head.
But I think that a strange person is turned away than everyone.
But I think that a strange person living
a happy life and a interesting.
General people always are not interesting in generally.
A strange person can whatever don't worry the public notice.
I think that a strange person have a courage.
But a strange person is sent a look of prejudice by a people.
I also want to a courage and power of a strange person.
I like a strange person very much.

— Mika Imae

650

They are really strange guys. They love saké, Tobaco, women
and other all bad things. They love break the regulation.
And They don't have few money.
Guys always say only something like
dream. They are the worst person I had ever seen.
But I love them. My bad company was formed by six person.
"A really important thing can't see,"
said Antoine de Saint-Exupéry.
I was teached it by them.
When I talk to them about UFO, God, Stars, nature,
make love . . . etc. I'm very very happy. I promised that
we'll fly in the sky someday.
And they are pure in the world, They are kind to their friends.
They are the important person by me. They are my all.
I've been on top of the world since I met them.

— Naomi Oue

Animals Has A Right!

People does a foolish!
My neighborhood took dog home and reared it,
but he took his dog to the Health Center.
For the reason that the dog didn't warm up to him.
Do you know the Health Center?
The Health Center kill dogs or cats and is
carried out an experiment on animals.
People lack all sense of responsibility oughtn't to
rear animals.
Animals isn't toy!
Animals feeds on food, walks, sleeps too.
People is willfulness!
I don't rear animal, because I don't want it
present at a animal deathbed.
Animals has a right!
People make a mistake, but want to notice
this mistake earlier.
It is cruel to kill animals.
I think Animals would like to talk a walk and
have to do with.
I have to keep on this movement!

— Tomoko Morita

651

One day I'm playing an hourglass in the park and
I was run after the dog. I sprain my leg. I was
bitten it. I'm dripping blood from my leg.
I run about in there.
This time I think it. I can't overtake. So
when I was walking a return road, huge dog stood
in front of me. the dog run after me, too.
So the dog make a counter attack me.
The dog stood a pair of legs and he can run
100 meters in 9 seconds.
Then I sweat at my woke up.
My leg's sprain was bitten by a mosquito mark.

— Kumi Kamada

I like get some exercise. One day, I climbed mountain with
my cousin's dog. Frankly speaking, I don't like my cousin's
dog. Because, cheekly and not pretty, glutton. I was work
by a dog. Everyday, I have a habit of take a walk. A dog
had the impudence to buy feed. By the way, I like
cooking. But, I cook feed for a dog, a dog seem to not
taste good. At last, I was gone a dog. Because the
feed is wasteful, I sold the food cheap. My cousin visited
my home. She said, "Thanks, take care of my dog."
After a few weeks, I was strike by my cousin. My face
was expand bigly.

— Yumiko Shigeoka

Very an Unpleasant Dream Was

It's nearly. When I was sleeping, I heard strange noise so
wake up. And as terrible. I was kept laid. And
I found that the strange noise was spoke by somebody.
The strange noise is the louder around my ears.
I think that thief was enter my room. But no one stay
here. Just only we. Who's noise? Maybe ghost.
Because then I, wasn't move my body. I was surprised
at this accident. But it's true. when I wasn't move
my body, my dog was sleeping next to me. But my dog
was no happen.
Next day I told my mother "I was terrible accident last night."

But my mother unbelieved such a accident. But I seemed that
I never forget this accident.
I don't know why is ghost is chosen me.
It's a terrible . . .

— Aya Iwaki

I remember a dream clearly. I lived along by a lake.
The lake is rumored that if somebody drowned in the lake,
three bodies went down to the bottom of the lake, will
never discovered, and a part of the bottom of the lake
is an icy bed, so these drowned bodies will not go bad
permanently, like a died in a snowfield person is discovered
with alike alive.
One day, I went to fishing with my neighborhoods.
(some are alive now, but others are died) at lake
I wasn't fish the lake, instead of fish, I fished a
person in the lake, not a drowned but a alive now I know
my wife, my children, my friend's and my dear dog!!
I fished an alive person in the lake, instead of my neighbor-
hoods who are died now, went down to the bottom of the
lake.
At last I fished my wife, and I looked back. There's my
daddy. he walked to the lake. I didn't say anything. I only saw
his back. on the way to the lake he turned me and say
"Good luck!"
After I awaken I'm just a little cry, in my bed.

— Midori Kawamura

I had an unpleasant dream a few days before.
I was taking as usual lessons. when the teacher
began to call role, I already took my seat.
When I was called my name, I gave a definite answer.
For all that, my answer
seemed to be heard nothing him.
So, I said to him, "I'm here." For all that, the teacher
ignored me and began to lesson. I said many a time,
but impossible. I was only here, and I was not
no one anything to do at all.
Very an unpleasant dream was.
Maybe it will the dream came true.

— Minako Nagamatsu

I was very beautiful and a very rich
all over the world.
Therefore I was envied by everyone.
But my character was one of a fault.
One evening, angel said "You don't yourself
until now from tomorrow."
The next day, I changed myself like dream
when I got up in the morning. In short,
I changed a very poor and very ugly.
My lovely fiancé went away from me.
I didn't easily recover.
It was a good chance, yourself could
understand. Therefore I took part in a
volunteer circle. One day I met a lovely
age there. He was kind and encouraged
when I was depressed.
I loved more and more him.
He so just as admitted. I was very glad
and very happy. He says "You have
no connection your looks. In sort, more important
human hurt." I could forgot old fiancé
thanks to him.
And one day I went to his home with
homemade a lunch.
To my surprise, he married and have
children. In short I was deceived by him.
Then I couldn't believe everything.
Then I became morbid fear.

— Yuko Abe

Doubt His Dead

My grandfather died in 1979. He liked bread.
So I sometimes bought bread near my house.
But I didn't remember him in the least.
He died when I probably five years old.
So I don't know his face.
My mother told me about him. He was kind for me.
I told about him by her. It was great shock to me.
I was present at a my grandfather's deathbed. I stayed
at there alone. But I didn't remember it at all.
When I gave him a piece of bread, he died.

I doubt his dead. I was very shock. But my mother told me.
She said, "You are not responsible for his dead."
I believed her. I like my grandfather.

— Maki Nakakoge

When I was sixteen, he was killed in traffic accident.
He was his dearest man for me.
I have never been faced death.
At first, I could hardly believe such thing.
but, It was a plain fact.
At that time, I rode with him on his motorbike,
but he was only died,
when I was just before my eyes the sight,
I cried his name (all my might and main) repeating.

— Naomi Furuhara

I have many relatives, so I attend the funeral well.
the year before last my grandfather's cousin died too.
I wounder how many people died. I never want someone to die anymore. I have a grandfather. He kind me.
I sometimes visit his house. he felt well when I visited before.
However he is ill now. I heard he has been in hospital.
I thought he might catch a cold. because He aggravates his disease well. But he has been cerebral thrombosis.
I was surprised.
and I went to the hospital in a hurry,
he was lieing on the bed as if he were not.
All relatives visited him. For the first time,
I'm thinking about person's death.
I wish grandfather can move in good again.
I don't want remembering someone who died.

— Kazumi Nobori

Talk My Mind

A neat man moved into next to my house a month ago.
He seemed a company man.
He usually worked at the company the daytime, but he seemed to

something doing at his house every evening.
And I checked up his belongings.
Then I know he was a new half.
He maked up and wear a dress and was fascinated a looking mirror. when I happened to pass at his house one evening.
Therefore he had a taste it every evening.
He escalated every evening. Because He used to went out of his house for meet someone.
Of course, I followed him a very interesting.
To be surprise, He met my father.
I was shocked it. Next day, I told my father "Your goddess true colors is a mysterious neighbor."
And my father shocked it as good as me.

— Yuko Abe

It's the most important for all my life that is to meet with a lot of people. I've thought much of that recently. to tell the truth I'm egoistic but it was stronger before I entered the college. I wasn't anxious for no one what one's think about my action but I met various kind of person at this college, then I'm felt this self-satisfaction. Till now Even I feel satisfaction I doesn't matter anyone else. Because I had thought I've a strong mind. Even if I'd tackle something bad thing I can conquer within someones help directly.
But I didn't. I had been worried for a long time. and I had lost confidence in myself.
I didn't know what to do when I wanted to cry because I didn't understand about myself. my friends took me out and they told me their personal worry.
I had never felt what the most important friends are.
However, I haven't told yet my mind for friends actually. as I haven't ever told my mind for anyone,
I can't talk my mind well.
Now I always listen friends talking and
I'm learning for friends how to talk directly but they often says "I don't understand what you want to say."
I think that I have to do now is more talking with a lot of people of around me than before. then I can stop to being worried. I feel to talk my thought briefly is one of condition for grown up to be a woman.

— Keiko Sakae

Notes

1. Wittgenstein, L. (1953). *Philosophical investigations.* (London: Basil Blackwell). This quotation can be found in Part 1, Section 6.

2. Jean-Jacques Lecercle (1985), a French professor of English, has elaborated on this point in his insightful book *Philosophy through the looking glass.* (La Salle, IL: Open Court):

   > *As a teacher of English, of necessity imposing on my students the standard dialect, I know that I am passing on a version of the language that most English speakers would laugh at.*

3. Perhaps the dangers are more real than I had supposed. In October, 1992, the Japanese high school exchange student, Yoshihiro Hattori, was shot and killed in Baton Rouge, Louisiana, in part, it is assumed, because he failed to understand the command *"Freeze!"*

4. In 1991, I published the essay, "The Accidental Emergence of the Para-Poetic," in the *Bulletin of the Suzugamine Women's College Cultural and Social Science Studies, 38,* Hiroshima, Japan. In that essay I introduced the concept of the para-poetic and attempted to provide an historical context for looking at student writings from a literary perspective.

5. The notion of language speaking has been elaborated on and developed by Martin Heidegger (1971):

   > *Man acts as though he were the shaper and master of language, while in fact language remains the master of man . . . for strictly, it is language that speaks. Man first speaks when, and only when, he responds to language by listening to its appeal.*

   From "Poetically man dwells . . ." in A. Hofstadter (Trans.), *Poetry, language, thought* (New York: Harper & Row), pp. 215-216.

6. The student Yoshiko Iwata made these comments. Due to her excellent but incomplete English, conversations with her have proven valuable in exploring some of the smaller, more subtle errors in English and their effects in conveying (or preventing) a meaning or a point of view.

7. This idea is expanded further with the discussion of "container metaphors" in George Lakoff and Mark Johnson (1980), *Metaphors we live by* (Chicago: The University of Chicago Press), pp. 29-30.

8. In Jean-Jacques Lecercle's (1990) *The violence of language* (London:

Routledge), he speaks of the same phenomenon of language's excess and refers to it as "the remainder:"

> *Language is no longer a mere instrument, it seems to have acquired a life of its own. Language speaks, it follows its own rhythm, its own partial coherence, it proliferates in apparent, sometimes violent, chaos. . . . There is another side to language, one that escapes the linguist's attention, not because of his temporary failure or failings, but for necessary reasons. This dark side emerges in nonsensical and poetic texts, in the illumination of mystics and the delirium of logophiliacs or mental patients. . . . I have called it "the remainder"* (pp. 5-6).

9. The concept of dream condensation was developed by Sigmund Freud and discussed in his book *The interpretation of dreams.* (1965, New York: Basic Books), pp. 312-339.

Author

CLARK LUNBERRY, Lecturer, Department of English, Suzugamine Women's College, 4-6-18 Inokuchi, Nishi-ku, Hiroshima 733, Japan. Interests: linguistics and literature, second language acquisition.

658

Calendar of Events *

1994

Global Awareness Society International, 3-5 June 1994, Chicago. *Information*: Marvin Harris, English Department, East Texas Baptist University, Marshall, TX 75670-1498.

American Classical League, 22-25 June, Oxford, OH. *Information*: Harry C. Rutledge, ACL, University of Tennessee, Knoxville, TN 37996.

Society for the History of Authorship, Reading and Publishing, 14-16 July, Washington. *Information*: John Y. Cole, Center for the Book, Library of Congress, Washington, DC 20504; 202/707-5221.

Computers in Applied Linguistics, 9-13 July, Ames. *Information*: Carol Chapelle, 203 Ross Hall, Iowa State University, Ames, IA 50011; e-mail <carolc@iastate.edu>

American Association of Teachers of French, 17-20 July, Quebec City. *Information*: AATF, 57 E. Armory Ave., Champaign, IL 61820.

International Association for Southeast European Studies, September date to be announced, Greece. *Information*: AIESEE, Department of Slavic Languages, CB #3165, Dey Hall, University of North Carolina, Chapel Hill, NC 27599-3165.

German Studies Association, 29 September-2 October, Dallas. *Information*: Diet Prowe, Department of History, Carleton College, Northfield, MN 55057; 507/663-4216; fax 507/663-4204; e-mail <dprowe@carelton.edu>

Pennsylvania State Modern Language Association with Modern Language Association of Philadelphia and Vicinity, 30 September-1 October, Philadelphia. *Information*: Diana Regan, 411 Dartmouth Rd., Bryn Mawr, PA 19010.

Second Language Research Forum, 6-9 October, Montreal. *Information*: Victoria Murphy, Faculty of Education, McGill University, 3700 McTavish St., Montreal, PQ, Canada H3A 1Y2; 514/398-6982; fax 514/398-4679; e-mail <czvm@musica.mcgill.ca>

Confederation of Oregon Foreign Language Teachers, 7-8 October, Sun River. *Information*: Chela Kocks, Southern Oregon State College, Ashland, OR 97520.

* Courtesy of *The Modern Language Journal* *(University of Wisconsin Press)*

Pennsylvania State Modern Language Association, 7-8 October, State College. *Information*: Eileen Glisan, Indiana University of Pennsylvania, 472 Sutton Hall, Indiana, PA 15705.

Japan Association of Language Teachers, 8-11 October, Ehime. *Information*: JALT, Glorious Tokyo #301, 2-32-10 Nishinippori, Arakawa-ku Tokyo 116, Japan; tel. +81-3-3802-7121; fax +81-3-3802-7122.

American Translators Association, 12-16 October, Austin. *Information*: ATA, 1735 Jefferson Davis Highway, Suite 903, Arlington, VA 22202-3413; 703/892-1500; fax 703/892-1501.

Canadian Parents for French, 13-15 October, Fredericton, New Brunswick, Canada. *Information:* W. Green, CPF National Office, 309 Cooper Street, Suite 210, Ottawa, ON K2P 0G5; 613/235-1481, fax 613/230-5940.

Foreign Language Association of North Carolina, 13-15 October, Greensboro. *Information*:: Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; 919/763-4009.

Washington Association of Foreign Language Teachers, 14-15 October, Wenatchee. *Information*: Molly King, 15341 SE 182nd St., Renton, WA 98058.

Foreign Language Association of North Dakota, 19-21 October, Minot. *Information:* Herbert Boswau, PO Box 8198, Grand Forks, ND 58202-8198.

Minnesota Council on the Teaching of Foreign Languages, 20-21 October, location to be announced. *Information*: Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

Classical Association of the Empire State, 20-22 October, Syracuse. *Information*: Jill Crooker, 15 Sutherland St., Pittsford, NY 14534.

Texas Foreign Language Association, 20-22 October, Dallas. *Information*: Eugenia Simons, 1320 Modiste Dr., Houston, TX 77055.

New York State Association of Foreign Language Teachers, 21-23 October, Kiamesha Lake. *Information*: Helene Combopiano, 71 Grand Blvd., Binghamton, NY 13905.

ndiana Foreign Language Teachers Association, 27-28 October, ndianapolis. *Information*: Carol E. Clemens, 7823 White Dove Court, ndianapolis, IN 46256.

owa Foreign Language Association, 28-29 October, Des Moines. *Information*: Paul Hoekstra, Bureau of Curriculum & Instruction, Grimes tate Office Building, Des Moines, IA 50319.

Nebraska Foreign Language Association, 28-29 October, Omaha. *Information*: Susan Smith, 1515 West Monor Dr., Lincoln, NE 68506.

Massachusetts Foreign Language Association, 4-5 November, Burlington. *Information*: Georg Steinmeyer, Black Mountain Rd., RR1, Box 549, Brattleboro, VT 05301.

National Association of Self-Instructional Language Programs, 4-5 November, Washington. *Information*: NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19922.

Wisconsin Association of Foreign Language Teachers, 4-5 November, Appleton. *Information*: William W. Kean, Suring Public Schools, Box 158, Suring, WI 54174.

American Association for the Advancement of Slavic Studies, 17-20 November, Philadelphia. *Information*: AAASS, Stanford University, 128 Encina Commons, Stanford, CA 94305-6029.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 18-20 November, Atlanta. *Information*: ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; 914/963-8830; fax 914/963-1275.

Modern Language Association of America, 27-30 December, San Diego. *Information*: MLA, 10 Astor Place, New York, NY 10003-6981; fax 212/477-9863.

1995

Texas Foreign Language Association, 21-22 February, South Padre Island. *Information*: Eugenia Simons, 1320 Modiste Dr., Houston, TX 77055.

National Association of Independent Schools, 1-4 March, Boston. *Information*: Claire Whalen, NAIS, 1620 L St. NW, Washington, DC 20036.

Southern Conference on Language Teaching, 2-4 March, Charleston. *Information*: SCOLT, Modern Foreign Languages, Valdosta State University, Valdosta, GA 31698.

Teachers of English to Speakers of Other Languages, 28 March-1 April, Long Beach. *Information*: TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2705.

Medieval Academy of America, 30 March-1 April, Boston. *Information*: Medieval Academy of America, 1430 Massachusetts Ave., Cambridge, MA 02138.

Central States Conference on the Teaching of Foreign Languages with Pacific Northwest Conference on Foreign Languages, Southwest Conference on Language Teaching, and Colorado Congress of Foreign Language Teachers, 30 March-2 April, Denver. *Information*: CSCTFL, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; 608/246-6573; fax 608/246-6880.

Georgetown University Round Table, 2-5 April, Washington. *Information*: School of Languages & Linguistics, 303 Intercultural Center, Georgetown University, Washington, DC 20057-1067.

Association for Asian Studies, 6-9 April, Washington. *Information*: Association for Asian Studies, 1 Lane Hall, University of Michigan, Ann Arbor, MI 48109.

Northeast Conference on the Teaching of Foreign Languages, 6-9 April, New York. *Information*: Northeast Conference, St. Michael's College, 29 Ethan Allen Ave., Colchester, VT 05446.

California Foreign Language Teachers Association, 28-30 April, Los Angeles. *Information*: Hal Wingard, 6450 Lance Way, San Diego, CA 92120.

International Association for Learning Laboratories, 23-28 May, South Bend. *Information*: IALL, University of Pittsburgh, LAI G-47 C.L., Pittsburgh, PA 15260.

American Association of Teachers of French, July dates to be announced, San Antonio. *Information*: AATF, 57 E. Armory Ave., Champaign, IL 61820.

American Association of Teachers of German with **Internationaler Deutschlehrerverband**, 4-7 August, Stanford University. *Information*: AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; 609/795-5553; fax 609/795-9398.

International Council for Central and East European Studies, 6-11 August, Warsaw. *Information*: George Fowler, Slavic Department, Indiana University, Bloomington, IN 47405; fax 812/855-2107; e-mail <gfowler@indiana.edu>

Iowa Foreign Language Association, October dates to be announced, Des Moines. *Information*: Paul Hoekstra, Bureau of Curriculum & Instruction, Grimes State Office Building, Des Moines, IA 50319.

Pennsylvania Modern Language Association, 6-7 October, Pittsburgh. *Information*: John Pierce, Muhlenberg College, Foreign Languages and Literatures, Allentown, PA 18104-5586.

Foreign Language Association of North Carolina, 12-14 October,

Greensboro. *Information*: Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; 919/763-4009.

Foreign Language Association of North Dakota, 18-20 October, Fargo. *Information*: Herbert Boswau, PO Box 8198, Grand Forks, ND 58202-8198.

Minnesota Council on the Teaching of Foreign Languages, 19-20 October, location to be announced. *Information*: Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

Classical Association of the Empire State, 19-21 October, Syracuse. *Information*: Jill Crooker, 15 Sutherland St., Pittsford, NY 14534.

New York State Association of Foreign Language Teachers, 20-22 October, Kiamesha Lake. *Information*: Helene Combopiano, 71 Grand Blvd., Binghamton, NY 13905.

Texas Foreign Language Association, 26-28 October, Austin. *Information*: Eugenia Simons, 1320 Modiste Dr., Houston, TX 77055.

American Association for the Advancement of Slavic Studies, 26-29 October, Washington. *Information*: AAASS, Stanford University, Jordan Quad/Acacia Bldg., Stanford, CA 94305-4130.

Japan Association of Language Teachers, 2-5 November, location to be announced. *Information*: JALT, Glorious Tokyo #301, 2-32-10 Nishinippori, Arakawa-ku Tokyo 116, Japan; +81-3-3802-7121; fax +81-3-3802-7122.

National Association of Self-Instructional Language Programs, 3-4 November, Washington. *Information*: NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19122.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 18-20 November, Anaheim. *Information*: ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; 914/963-8830; fax 914/963-1275.

American Association of Teachers of Slavic & East. European Languages & Literatures, 27-30 December, Chicago. *Information*: AATSEEL, Russian & Slavic Languages, University of Arizona ML 340, Tucson, AZ 85721.

Modern Language Association of America, 27-30 December, Chicago. *Information*: MLA, 10 Astor Place, New York, NY 10003-6981; fax 212/477-9863.

1996

National Association of Independent Schools, 5-8 March, Washington. *Information*: Claire Whalen, NAIS, 1620 L St., NW, Washington, DC 20036.

Teachers of English to Speakers of Other Languages, 26-30 March, Chicago. *Information*: TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2705.

Central States Conference on the Teaching of Foreign Languages, 28-31 March, Louisville. *Information*: CSCTFL, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; 608/246-6573; fax 608/246-6880.

Georgetown University Round Table, 31 March-3 April, Washington. *Information*: School of Languages & Linguistics, 303 Intercultural Center, Georgetown University, Washington, DC 20057-1067.

Association for Asian Studies, 11-14 April, Honolulu. *Information*: Association for Asian Studies, 1 Lane Hall, University of Michigan, Ann Arbor, MI 48109.

Northeast Conference on the Teaching of Foreign Languages, 18-21 April, New York. *Information*: Northeast Conference, St. Michael's College, 29 Ethan Allen Ave., Colchester, VT 05446.

California Foreign Language Teachers Association, 26-28 April, Sacramento. *Information*: Hal Wingard, 6450 Lance Way, San Diego, CA 92120.

International Association of Applied Linguistics, 4-9 August, Jyväskylä, Finland. *Information*: Udo O. H. Jung, Language Centre, University of Bayreuth, D-95440 Bayreuth, Germany; (0921) 553595; fax (0921) 553594; e-mail <udo.jung@uni-bayreuth.de>

Minnesota Council on the Teaching of Foreign Languages, 17-18 October, location to be announced. *Information*: Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

New York State Association of Foreign Language Teachers, 18-2-October, Kiamesha Lake. *Information*: Helene Combopiano, 71 Grand Blvd., Binghamton, NY 13905.

Texas Foreign Language Association, 18-20 October, Houston. *Information*: Eugenia Simons, 1320 Modiste Dr., Houston, TX 77055.

Foreign Language Association of North Dakota, 23-25 October,

Bismarck. *Information*: Herbert Boswau, PO Box 8198, Grand Forks, ND 58202-8198.

Foreign Language Association of North Carolina, 24-26 October, Greensboro. *Information*: Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; 919/763-4009.

National Association of Self-Instructional Language Programs, 1-2 November, Washington. *Information*: NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19122.

American Council on the Teaching of Foreign Languages with **American Association of Teachers of German**, 22-24 November, Philadelphia. *Information*: ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; 914/963-8830; fax 914/963-1275.

Modern Language Association of America, 27-30 December, location to be announced. *Information*: MLA, 10 Astor Place, New York, NY 10003-6981; fax 212/477-9863.

1997

Northeast Conference on the Teaching of Foreign Languages, 3-6 April, New York. *Information*: Northeast Conference, St. Michael's College, 29 Ethan Allen Ave., Colchester, VT 05446.

Foreign Language Association of North Carolina, 10-12 October, Greensboro. *Information* Wayne Figart, 204 N. 16th St., Wilmington NC 28401; 919/763-4009.

Modern Language Association of America, 27-30 December, location to be announced. *Information*: MLA, 10 Astor Place, New York, NY 10003-6981; fax 212/477-9863.

Associations, please note: Apprise us as soon as your conference dates and locations have been confirmed. We shall then include this information in the *Calendar of Events* section of *Applied Language Learning*.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning (ALL)* is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editor encourages the submission of original manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006.

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (3rd Ed., 1983) available from the American Psychological Association, 1200 17th Street, NW, Washington, DC 20036.

Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating:

full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

Jane C. Doe, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

The list of references should be submitted on a separate page of the manuscript with the centered heading: *References*. The entries should be arranged alphabetically by surnames of authors. The sample list of references below illustrates format for bibliographic entries:

Dulay, H. & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 16 (1)*, 93-95.

Harris, D. P. (1969). *Testing English as a second language.* New York: McGraw-Hill.

Reference citations in the text of the manuscript should include the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the material that is being quoted originally appeared, e.g., (Jones, 1982, pp. 235-238). All works cited in the manuscript must appear in the list of references, and conversely, all works included in the list of references must be cited in the manuscript.

Notes should be used for substantive information only, and they should be numbered serially throughout the manuscript. Subsequently they all should be listed on a separate page titled *Notes*.

Specifications for Floppy Disks

Where feasible, manuscripts are preferred on disk (DOS or XEROX Global View). Manuscripts produced on one of the DOS-based or Macintosh systems should be formatted as *MS-DOS* file at double density and entered on the disk as an ASCII file if possible.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____
2. Saved as an ASCII file? YES_____NO _____
3. Disk is formatted as: double: _____high density_____
4. Remarks:

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

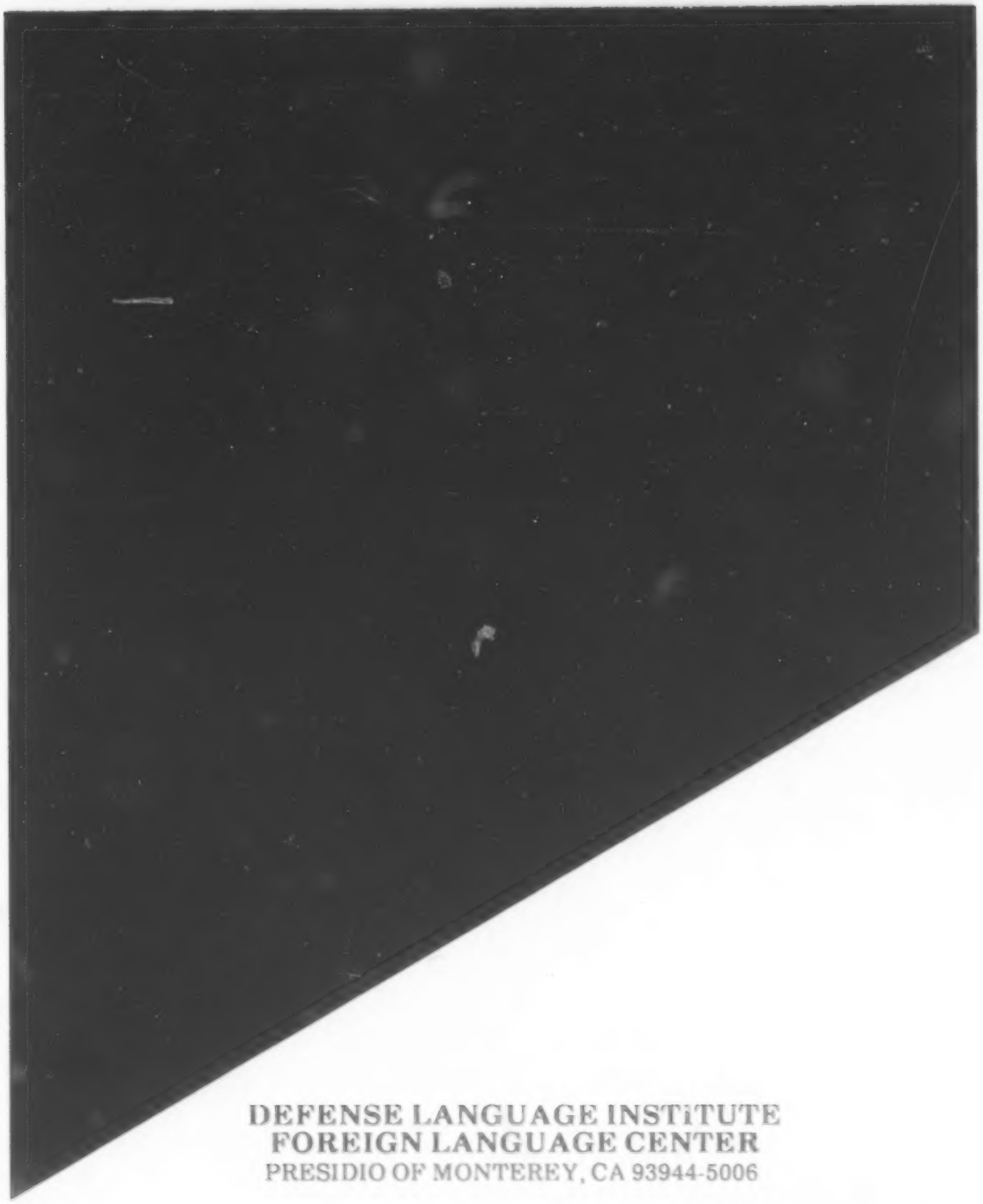
Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

668

* U.S. GOVERNMENT PRINTING OFFICE: 1994-583-000/11024

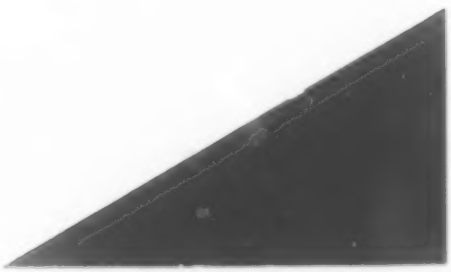
PIN: 072639-UN-000

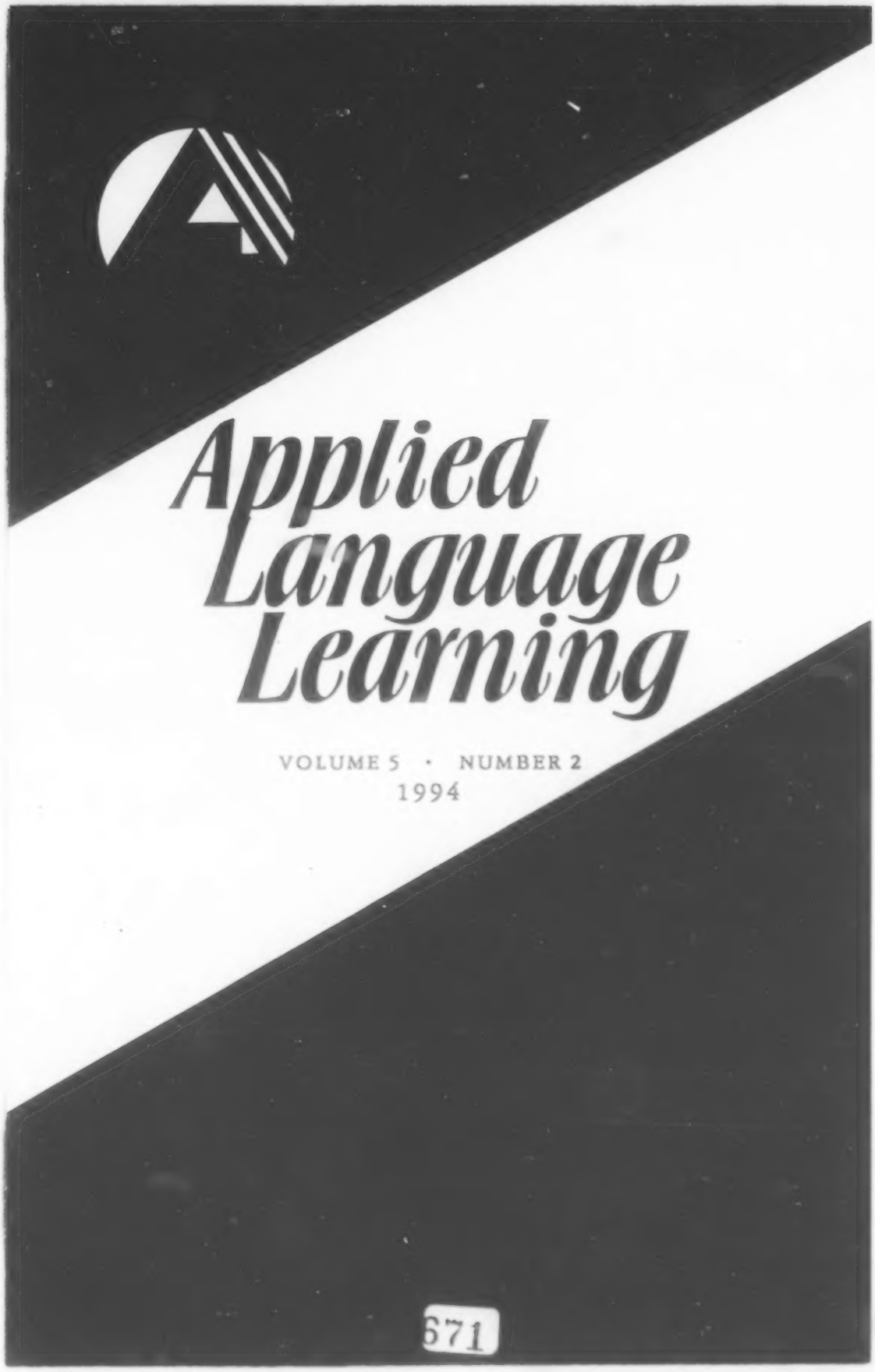


DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

670

Professional Bulletin 65-94-1
United States Army
Approved for public release. Distribution is unlimited.





Applied Language Learning

VOLUME 5 · NUMBER 2
1994

671

Applied Language Learning

Volume 5 • Numbers 1 & 2

1994

Defense Language Institute
Foreign Language Center
Presidio of Monterey

672

Applied Language Learning
PB 65-94-2

The mission of Professional Bulletin 65-94, *Applied Language Learning* (US ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

*Applied Language Lea*rning, published semi-annually by the Defense Language Institute, Foreign Language Center and Presidio of Monterey, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

Official:

MILTON H. HAMILTON
Administrative Assistant to the Secretary of the Army
07762

GORDON R. SULLIVAN
General, United States Army
Chief of Staff

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to DLIFLC. *Applied Language Learning* is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster:

Send change-of-address information to:

Applied Language Learning
Office of the Provost
Defense Language Institute, Foreign Language Center, and Presidio of Monterey
Monterey,CA 93944-5006
Telephone: 408 242-5638
DSN: 878-5638

Applied Language Learning

PB-65-94-2
USSN 1041-6790

Lidia Woytak
Editor

Editorial Board

Kathleen M. Bailey	Professor, MIIS
David P. Benseler	Professor, Case Western Reserve University
Christine Campbell	Project Director, DLI
John L. D. Clark	Dean, DLI
Ray T. Clifford	Provost, DLI
Martha Herzog	Asssociate Provost, DLI
Stephen D. Krashen	Professor, USC
John A. Lett, Jr.	Director, Research, DLI
June K. Phillips	Dean, Weber State University
Leo van Lier	Professor, MIIS

DLI Staff

Editorial Coordinator, Christa Rutsche
Software Consultant, Milan Herben
Graphics, Steven Jackson
Wordprocessor, Pricilla Coffman

A Note of Appreciation

The individuals listed below served as referees during the preparation of Volume 5 of *Applied Language Learning.* The editor and the editorial board express their gratitude for expert service to:

Kathleen M. Bailey
Christine Campbell
Rod Ellis
Martha Herzog
Eli Hinkel
Gordon Jackson
Steven Krashen
B. Kumaravadivelu
John A. Lett, Jr.
Rebecca L. Oxford
Howard Roland
June K. Phillips
Victor M. Shaw
John Thain
Leo van Lier

Applied Language Learning

1994 Volume 5 Number 2

Articles

1 Curricular Innovation: Issues and Problems
Numa Markee

31 Computer-Assisted Language Learning in Cooperative Learning
Lydie E. Meunier

57 FL/L2 Reading, Grammatical Information, and the Input-to-Intake Phenomenon
David J. Shook

Research Report

95 Reading versus Grammar: What Students Think Is Pleasureable and Beneficial for Language Acquisition
Jeff McQuillan

General Information

101 Calendar of Events

107 Information for Contributors

Applied Language Learning
1994, Vol. 5, No.2, pp. 1-30

ARTICLES

Curricular Innovation: Issues and Problems

Numa Markee
University of Illinois at Urbana-Champaign

Although considerable progress has been made in understanding how to design *innovative syllabuses, much less is known about the issues and problems that must be solved when* implementing *such syllabuses. To this end, curricular innovation is defined as a phenomenon that involves managing developmental change in the design, implementation, and maintenance of teaching (and/or testing) materials, methodological skills, and pedagogical values that are perceived as new by individuals who comprise a formal (language) education system. Based on this definition, it is proposed that three dimensions of educational change that are referred to as an "innovation triangle" – namely, the development and use by teachers of new materials, methodological skills, and pedagogical values - constitute the core of all curricular innovation. Using a project in curricular innovation at the University of Illinois at Urbana-Champaign (UIUC) as a case study, this paper first reviews in what sense a task-based approach to syllabus design constitutes an innovation. Then, the UIUC model of curricular innovation is briefly outlined. Next, the innovation triangle is used as a framework for analyzing the extent to which a complete and coherent process of innovation has been implemented in the UIUC project. The paper concludes by suggesting that the implementation of task-based syllabuses by ESL (English as a Second Language) teachers at UIUC has been marked by uneven development along the three dimensions of change identified above. The paper ends with some general considerations about the nature of curricular innovation.*

Introduction

The last two decades have witnessed the development of a number of innovative syllabuses that have been of great interest to second language curriculum specialists and teachers. For example, the notional-functional syllabus (Wilkins, 1976), the

©1994, Numa Markee

process syllabus (Breen & Candlin, 1980; Breen, 1984; 1987; Candlin, 1984; 1987), the Natural Approach (Krashen & Terrell, 1983), the procedural syllabus (Prabhu, 1987) and task-based language teaching (TBLT) Kumaravadivelu, 1991, 1993; Long, 1985, 1989, 1991, 1993; Long & Crookes, 1992; 1993; Loschky & Bley-Vroman, 1993; Nunan, 1989, 1992, 1993; Pica et al., 1993) have all significantly changed the way language teaching professionals think about designing language education programs.

This work has undoubtedly deepened the language teaching profession's understanding of issues related to the *design* of innovative language syllabuses. However, the profession's understanding of the issues that pertain to the *implementation* of such innovations is still under-developed. Since the ultimate proving ground of all pedagogical innovations is the classroom, language teaching professionals need a framework that helps them address implementation problems in a fashion that is both pragmatic and theoretically motivated. To this end, I propose that curricular innovation involves managing developmental change in the design, implementation and maintenance of *teaching (and/or testing) materials*, *methodological skills* and *pedagogical values* that are perceived as new by individuals who comprise a formal (language) education system.

I do not intend to provide a comprehensive discussion of all the constituent elements of this definition here (see Markee [1993] for such an overview). For present purposes, let me concentrate on discussing the theoretical and practical implications of those parts of the definition that are *underlined*, which constitute the core of any attempt at educational innovation. In this regard, it has been suggested that the development and use of new materials by teachers may be the easiest kind of change to promote. This is because materials are the most tangible products of any process of educational innovation (Fullan, 1982a, 1982b, 1993). At the same time, we should realize that "the relationship between behavioral change (e.g., teaching approach) and changes in beliefs is complicated. Logically, one might think of beliefs changing first, which, in turn, lead to new behavior associated with the belief. Practically, however, there is considerable socio—psychological evidence to support the view that beliefs are learned through experience.." (Fullan, 1982b, p. 247).

The relationships among the three dimensions of educational change identified in the definition of curricular innovation are best conceptualized as an "innovation triangle," as shown in *Figure 1*. Here we see that a change in belief can precede a change in behavior; on the other hand, a change in behavior can lead to a change in values. The relationships among these dimensions are, in fact, so complex that, as Fullan states, "... perhaps it is sufficient for our purposes to recognize that the relationship between beliefs and behavior is reciprocal - trying new practices sometimes leads to questioning one's underlying beliefs; examining one's beliefs can lead to attempting new behavior" (Fullan, 1982b, p. 247).

The idea that the relationship between beliefs and behavior is reciprocal has obvious implications for teacher/curriculum development. If we accept the idea that teachers should "behave their way into new ideas and skills, not just think their way into them" (Fullan, 1993, p. 15), then this suggests that syllabus design/materials develop-

ment work by teachers can be used as a convenient entry point into the broader process of curricular innovation. But, as Fullan argues, the process of innovation is non-linear. Furthermore, it must also engage teachers in the more difficult business of developing their own methodological skills and changing their ideas about what constitutes good teaching. The problem, therefore, is that changes along one dimension of the innovation triangle are not necessarily accompanied by comparable changes along other dimensions of this triangle.

DEVELOPMENT AND USE OF NEW TEACHING MATERIALS
DEVELOPMENT AND USE OF NEW METHODOLOGICAL SKILLS
DEVELOPMENT OF NEW PEDAGOGICAL VALUES

FIGURE 1: *The dimensions of curricular innovation*

I use the innovation triangle posited above as a framework to analyze the extent to which change has been actually implemented in a curricular innovation project currently underway at the UIUC. More specifically, I show that, with respect to the first dimension of change identified above (changes in materials), ESL teaching assistants know how to develop innovative task-based language teaching units. When we investigate changes associated with the second dimension of change (changes in methodological skills), the situation becomes more complex. TAs know how to set up small group work, a procedure which is frequently used in TBLT. But their real-time methodological behaviors at particular decision points during small group work often result in classroom talk whose conversational structure is actually quite "traditional." And in terms of the third dimension of change (changes in pedagogical values), it seems that although most TAs are intellectually committed in general terms to TBLT, some of their specific beliefs concerning what constitutes good teaching are not necessarily compatible with this syllabus.

Thus, first, I review what TBLT is and show in what way it is an innovative approach to syllabus design. This review, which evaluates in some depth the kinds of changes that have been implemented by TAs in this project, is followed by a brief discussion of how the UIUC model of curriculum development is organized. Finally, I analyze the extent to which TAs have implemented change along all three dimensions of the innovation triangle. I conclude with some general insights about the nature of curricular innovation.

TBLT

It is widely accepted that language is a vehicle for communicating meaningful ideas and feelings that are important to learners. This concern with communication has led to at least three syllabus types that use tasks as the principal criterion for organizing language instruction: the procedural syllabus, the process syllabus, and TBLT. I do not propose to review the similarities and differences among these proposals here; this has already been done by other writers (see especially Long & Crookes, 1992, 1993). Rather, I use TBLT as an umbrella term that draws on ideas associated with the three syllabuses identified above.

These syllabuses are innovative for two different reasons. First, although TBLT syllabuses have objectively existed for some fifteen years now (see Prabhu, 1987), most TAs are unfamiliar with TBLT when they first start teaching ESL at UIUC. Thus, according to the definition of curricular innovation offered earlier, TBLT subjectively counts as a bona fide innovation for these particular individuals (Markee, 1993; A. Nicholls, 1983; Rogers, 1983).

Second, these syllabuses are innovative in terms of the evolution of second language teaching, because they are "analytic" rather than "synthetic" constructs (Wilkins, 1976). Synthetic and analytic approaches to syllabus design do not represent discrete categories of syllabus organization. In practice, all syllabuses may be placed on a synthetic-analytic continuum. However, in synthetic syllabuses, learning is organized in traditionally *linguistic* terms. Thus, the constituent linguistic elements of the target language (its syntax, semantics, lexis, morphology, phonology etc.) are pre-se-

lected and taught incrementally until the whole language is taught. Learners are exposed to small samples of language deliberately; their task is to "re-synthesize the language that has been broken down into a large number of small pieces with the aim of making this learning task easier" (Wilkins, 1976, p. 2). But in analytic syllabuses, learning is organized in behavioral terms of the social purposes that learners have for learning the target language. A focus on learners' needs leads to a focus on analyzing chunks of language performance. As Wilkins (1976:2) comments, "since we are inviting the learner, directly or indirectly, to recognize the linguistic components of the language behavior he is acquiring, we are in effect basing our approach on the learner's analytic capabilities."

There is broad agreement that TBLT represents an analytic approach to syllabus design. But there is also much disagreement concerning how to define tasks and TBLT. Furthermore, it is unclear what criteria should be used to select, grade, and sequence tasks. Thus, each of these issues in turn are addressed in this article.

What are Tasks?

Tasks consist of any "classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form" (Nunan, 1993, p. 59). Tasks require learners to produce chunks of language performance - that is, to communicate with the teacher and/or with each other - by using the linguistic resources they currently possess in the target language. As participants attempt to communicate, they modify their own and their interlocutors' speech to comprehend language which is beyond their current level of communicative competence. Learners thereby get the kind of negotiated comprehensible input which is thought to be a necessary (though not sufficient) pre-condition for second language development to occur (Long, 1985, 1989, 1991, 1993; Pica, 1987).

What is TBLT?

The analytic definition of task given above suggests that TBLT focuses as much on the process of learning as on its products. It also implies that TBLT is organized in terms of methodologically-induced opportunities to use and learn language. Thus, traditional distinctions between syllabus and methodology become blurred (Nunan, 1988). And in some forms of TBLT (i.e., the process syllabus), this "methodological syllabus" involves teachers and learners negotiating the content of instruction and/or the way in which classes are organized. Such negotiations can potentially change traditional classroom relationships between teachers and learners.

In the UIUC context, TBLT is defined as an approach to syllabus design in which chains of information-gathering, problem-solving, and evaluative tasks are used to organize language teaching and learning; these interdependent pedagogical tasks, which combine insights from sociolinguistic and psycholinguistic research, are designed methodologically to simulate the communicative events which learners encounter in second language-using environments.

What Criteria May be Used to Select, Grade, and Sequence Tasks?

The analytic definition of TBLT just given implies that task content may be selected on the basis of analyses of students' sociolinguistic needs (Long, 1985) and psycholinguistic wants (Brindley, 1984). Material may also be selected on the basis of its authenticity (Nunan, 1988) and sequenced according to the logical principle of task dependency (Johnson, 1982). This principle states that task one serves as input for task two, which in turn serves as input for task three, and so on. Empirical research also suggests that the length of turns, the number of elements a text contains, the genre to which a text belongs, the explicitness of a text's information structure, and the kinds of referents used to organize the discourse and participants' prior level of familiarity with the topic under discussion all have a bearing on the perceived difficulty of aural and/or written texts (Anderson & Lynch, 1988; Brown & Yule, 1983; Kiniry & Strenski, 1985; Nunan, 1985). These criteria may therefore be used to grade and sequence listening, speaking, reading, and writing tasks in TBLT.

Long, (1981) has demonstrated that some tasks are more difficult than others. For example, one-way information gaps, where only one participant holds new or unknown information, promote less obligatory conversational adjustments than two-way tasks, in which all participants hold information that is necessary to solve a problem. The research on task difficulty has led to the development of a task typology that may be used as a framework for choosing and using tasks (Pica et al., 1993). In addition, the use of different question types by teachers (i.e., referential versus display questions) has been shown to affect the syntactic complexity of students' responses (Brock, 1986; Nunan, 1987). Finally, Prabhu (1987) suggests that the amount and type of information that students must manipulate, the number of steps involved in completing cognitive operations, the accuracy with which students must perform and the extent to which they must draw on their knowledge of the world to solve a problem are all useful criteria for making, grading and sequencing decisions in TBLT.

The UIUC Model of Curriculum/Teacher Development

A fundamental assumption that underpins the UIUC project is that curriculum development is indivisible from teachers' professional development (Stenhouse, 1975). How is this assumption actually implemented? To enhance their professional development, TAs develop TBLT units in a pro-seminar which they take during their first semester of teaching. They then contribute these units to a central data bank, which all TAs can use in their ESL classes. The TAs' materials development activities are located in the theoretical model of curricular innovation that is depicted in *Figure 2*.

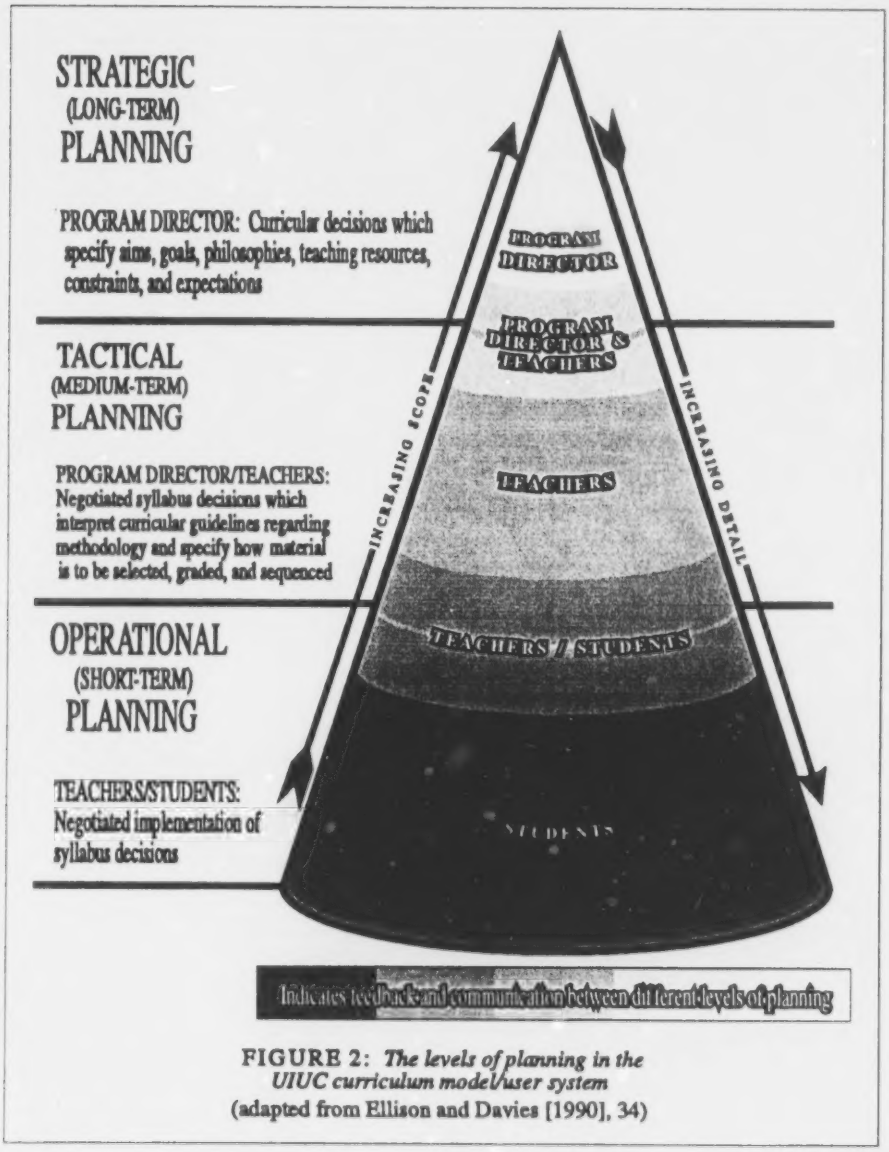


FIGURE 2: ***The levels of planning in the UIUC curriculum model/user system***
(adapted from Ellison and Davies [1990], 34)

In this model, three layers of planning are predicated. The *strategic* level of planning, which is equated with the establishment of general *curricular guidelines* for instruction, specifies a program' educational philosophy, outlines the purposes and content of instruction, identifies constraints on implementation and sets out the general criteria that are used for student and course evaluation. The *interpretation* of these guidelines is a *tactical* matter of *syllabus* planning, in which the content and/or the means of learning are negotiated between teachers and learners (Candlin, 1984).

683

BEST COPY AVAILABLE

And the "operational" level of planning (Ellison & Davies, 1990) subsumes lesson planning and teaching. Theoretically speaking, pedagogical innovation occurs as a synthesis of tensions between these three levels of planning (adapted from Candlin, 1984).

Figure 2 further specifies that the curricular or strategic level of planning involves long-term decision-making (Ellison & Davies, 1990); it also encompasses the entire program of ESL courses. This level of planning is the responsibility of the project director. Responsibility for the syllabus design or tactical level of planning is shared (i.e., negotiated) between the project director and the TAs; it involves medium-term decision-making, the scope of which encompasses parallel sections of a single course. Last, operational planning is the responsibility of TAs, who negotiate with their students how to implement the decisions made at the hierarchically higher level of syllabus design/tactical planning. This level of planning involves short-term, typically day-to-day, decision-making related to individual classes.

The model's funnel shape is borrowed from Tetenbaum and Mulkeen's (1986) model of teacher education, who wish to emphasize thereby that teacher development is an open-ended, developmental process. Since curriculum development and teacher development are viewed as two sides of the same coin, *Figure 2* underlines the fact that curriculum development is also an open-ended activity.

We may further conceptualize a curriculum as a *user system* in which the creative tension between different levels of planning is resolved by individuals who play socially-defined roles within a system. In the UIUC user system, the project director is a "change agent" or catalyst for curricular innovation; teachers are potential "adopters" (or indeed "resisters") of the proposed innovation (TBLT) and students are the "clients" for whom the innovation is ultimately developed (Kennedy, 1988). These roles are not mutually exclusive. In their role as syllabus designers and material developers, teachers act as change agents in their own right, while learners are potential adopters or resisters who may either accept or reject TBLT as a personally useful approach to language learning.

Let me now discuss the extent to which TAs have implemented changes associated with each one of the three dimensions of the innovation triangle identified earlier.

Dimension 1
The Development of New Teaching Materials

To date, TAs have developed a total of 55 TBLT units;[1] this number continues to grow each semester. As noted previously, units are written in the pro-seminar; they are then field-tested and revised in light of the feedback that TAs receive from peers in the pro-seminar, from their ESL students and from colleagues who agree to try these units out in their classes. Once the units have been revised, TAs contribute their work to the service courses' electronic banks of materials. The banks henceforth become a part of the pedagogical resources available for each course. From a materials development/use perspective, therefore, there can be no doubt that innovative TBLT

materials are being developed and used throughout the UIUC ESL service courses.

What do such materials look like and do they evidence an understanding by their authors of the principles of TBLT materials design discussed earlier in this paper? The extracts given below are from one of the best units produced to date (Broersma, 1993). Extract 1 gives an overview of the whole unit. The general types of class activities used in the unit's seven sections are summarized in Extract 2. And Extract 3 illustrates the specific tasks that are used in Section 2.

As shown by the overview in Extract 1, the unit aims to take learners (graduate students from a wide variety of disciplinary backgrounds) through the process of researching and writing an academic cause/effect essay, using the theme of archeological discoveries[2] as the carrier of content in the unit. The unit has the kind of behavioral orientation that characterizes analytic syllabuses in that it engages students in discovering how research is carried out and how knowledge is constructed. This focus on the process of research is highly relevant to the learners' intellectual development as graduate students who use English to do research at a US university.

Extract 1

The Iceman Speaks

Unit Overview

Section 1

Purposes

- To give students practice in recognizing and analyzing cause/effect relationships.
- To illustrate a multi-disciplinary perspective on problem-solving.
- To offer students opportunities to utilize various media as sources.
- To give practice in writing a scientific academic essay in cause/effect rhetorical mode.

Materials

NOVA video, *The Iceman.* Transcripts for "The Iceman" also available.

Fritz, S. (1993, February). Who was the Iceman? *Popular Science,* 46-49.

Wernick, R. (1988, March). What were Druids like, and Was Lindow Man one? (extract). *Smithsonian,* 146-147.

Houston, J. (1984, February). Sacred Slayings. *Science Digest,* 64-67, 102.

Green, T. (1972, June). Fifty years later, Tutankhamun's treasures are still a sensation. *Smithsonian,* 14-23.

Seidler, H., et al. (1992, October 16). Some anthropological aspects of the prehistoric Tyrolean Ice Man. *Science*, 455-457.

Marbach, W.D, and Phillips, J. (1987, November 2). New tools for an ancient dig. *Newsweek*, 80-81.

World Heritage Museum (Lincoln Hall), Brochure and Tour. Tours can be arranged in advance by calling #333-2360.

Supplemental Material

Warren, K. J. (1988). A philosophical perspective on the ethics and resolution of Cultural property issues. In P.M. Messenger (Ed.), *The Ethics of Collecting Cultural Property, Whose Culture? Whose Property?* (pp. 1-25) Albuquerque: University of New Mexico Press.

Sjovold, T. (1993, April). Frost and found. Natural History, 60-63.

Bohen, B. (1993). Videotaped lecture about the Iceman.

The Hatshepsut affair. (1989, May 13). *The Economist*, 49.

This overview also shows that the source materials used as input for the unit — primarily written texts from magazines and journals— are all authentic and intellectually-challenging. Also included are a videotape of a Nova program on the Iceman, and a videotaped lecture on this subject by a UIUC faculty member. As a class activity, students visit a campus cultural museum where they can ask questions on the issues raised in the source materials. In sum, the general focus on communication in an academic context and the selection of authentic texts and tasks that address the larger behavioral needs and interests of students are all consistent with the characteristics of TBLT outlined earlier in this paper.

The summary of activities shown in Extract 2 below shows in more detail how the process of carrying out research is broken down and logically organized and sequenced into various categories of pedagogical tasks. These categories include: students focusing on the thematic/rhetorical subject matter of the unit (Section 1); students gathering information from different sources and channels of communication in light of their own individual interests in the topic (Sections 2-4); students synthesizing information and presenting it in preliminary written and oral form (Section 5); students writing up first drafts of their research (Section 6); and finally, students peer-editing and re-writing these initial writing products in light of the feedback they receive from their teacher and from peers.

Extract 2
Summary of Activities

Section 1
√ *Introduction:* A brainstorming activity is followed by an analysis of cause/effect relationships.

Section 2
√ Information Gathering (Video): Students watch the NOVA video, "The Iceman." When the video is finished students freewrite and then break into interest sections and discuss cause/effect relationships in their group.

Section 3
√ Information Gathering: Students read articles about the Iceman, the Bog People, and King Tutankhamun with special attention to cause/effect information related to their interest sections. Students solve a problem or discuss a case study related to their interests.

Section 4
√ Information Gathering (Tour): In groups, students develop questions related to their interest sections to ask the tour guide in the World Heritage Museum (WHM). The second half of class is a tour in the WHM in Lincoln Hall.

Section 5
√ Synthesis: After working through the "abstract" handout, students use their notes from the WHM tour, transcripts of the NOVA program, and the articles already read to develop abstracts for their interest sections. After preparing individual abstracts, students present their abstracts to their group members for feedback.

Section 6
√ Output: Based on the abstracts, and utilizing sources, students write a cause/effect essay.

Section 7
√ Feedback: Using the peer editing handout as a guide, students exchange copies of their essays with students from other interest sections, and in groups of three or four have conferences about each paper.

√ Denotes worksheet(s) to go with activity.

Again, the organizing principles illustrated in Extract 2 are quite consistent with the characteristics of TBLT outlined previously in the paper. The design of this unit is on the analytic, not the synthetic end of the syllabus design spectrum. The only synthetic element in the design is the ultimate goal of getting students to write a cause/effect essay; however, most of the tasks which students must complete focus primarily on how to organize and develop such essays and only secondarily on the linguistic means used to write such an essay.[3] The result is a methodological syllabus of tasks, internally sequenced according to Johnson's (1982) logical principle of task dependency. In external grading/sequencing terms (i.e. in terms of when this unit is taught in relation to other units), it is taught toward the end of the semester because it focuses on a cognitively complex rhetorical style. Finally, by using a process-oriented, task-based methodology, the author creates a psycholinguistically-rich learning environment, which promotes SLA.

Finally, in order to understand better what TBLT materials actually look like, let us now look at Extract 3, Section 2, which consists of the first "information-gathering" activities used in the unit.

Extract 3
The Iceman Video
Information Gathering
Section 2

Materials

- Handout with pre- and post-video questions.
- NOVA video, *The Iceman*.
- Copies of portions of the transcript for *The Iceman*. (Optional)

Purpose

- To introduce the topic of the Iceman and other, similar, archaeological finds.
- To alert students to cause/effect relationships.
- To allow students to determine which aspects of the topic interest them most.

Instructions

Activity A

Reviewing
(5 minutes)

As a class, briefly review cause/effect relationships and the brainstorming activity from the previous class period.

Activity B

Watching the Video
(55 minutes)

Watch the NOVA program entitled "The Iceman." Try to discover cause/effect relationships in the video.

Activity C

Answering Questions
(10 minutes)

Answer the questions on the handout called "Questions about the Iceman Video" in as much detail as you can. When you finish answering the questions, think about which aspect of the Iceman discovery interests you the most. Which aspect would you most like to write a paper about? Here are some ideas:

Scientific

- the interaction of science with other fields;
- specific scientific applications, carbon-14 dating, for example;
- factors which can interfere with scientific analysis;
- how the scientific method is applied to archaeological finds

Historical/Cultural

- what the Iceman reveals about ancient cultures
- how the Iceman changes previously held beliefs about ancient people
- the processes historians use to analyze data
- limitations of the historical method

Political

- how political forces influence scientific research
- issues involved in the dispute over the Iceman between Austria and Italy. To whom does the Iceman really belong?

Activity D

Forming Interest Sections
(20 minutes)

> After determining which aspect of the topic interests you most, try to find three people who share your interest. Feel free to walk around the classroom to find others interested in the same topic. It is important that the groups consist of at least three people and no more than four. Also, try to find partners with different cultural backgrounds if possible. The teacher will make final decisions regarding the composition of groups if there are too many or not enough people with the same interests, or if there is not enough cross-cultural mixing in the groups. You may choose topics other than the ones listed above if you can find enough people who are interested in your topic.
>
> After the groups have been formed, in the time remaining, you should discuss within your group the cause/effect relationships mentioned in the video which had the most direct connection to the interest of your group. Keep a record of your discussion on your cause/effect handout.

Section 2 begins with a review (Activity A) of what occurred in the previous class; it also sets up the subsequent video-watching task (Activity B), in which students discover cause/effect relationships. As Nunan (1993:59) notes, tasks "should ... have a sense of completeness, being able to stand alone as ... communicative act[s] in [their] own right." These tasks not only exhibit this sense of individual completeness; the way in which one task flows into another also shows how the theoretically important principle of task dependency is actually implemented in TBLT.

Activity C sustains this principle further; students answer comprehension questions. Learners then move on to evaluate which aspect(s) of the Iceman discovery interests them most and decide on what topic they are going to write their unit papers. Suggested topics that students might explore include various aspects of the scientific, historical/cultural, or political dimensions of the discovery.

Students form groups in Activity D, where they use the topics of greatest personal interest to them as the basis for selecting the partners they will work with throughout the unit. Thus, the interest groups that are formed emerge primarily as a result of negotiations between students who, by finding out which other learners share their interests, have to bridge a complex information gap. In all these activities, the teacher often verbally paraphrases an activity's written instructions so that students can accomplish the task set as efficiently as possible.

As Clarke (1991) points out, ESL teachers in general often find that it is difficult to get learners to negotiate what a course will be about, since many students regard this as the teacher's job. Note, however, that Activities C and D get round this problem very neatly by setting up a real communicative need for learners to exchange information in

small groups. First, they must decide what aspect of the Iceman discovery interests them most; second, they must find other people who share similar interests; and third, they must identify with their partners which of the cause/effect relationships in the video might be relevant to their interest group. In short, Activities C and D illustrate a rather sophisticated use of jigsawed information-sharing tasks to negotiate the content of instruction and to determine who students will work with during the rest of the unit.

Dimension 2
The Development of New Methodological Skills

It is relatively easy to demonstrate that TAs are developing and using new TBLT materials. However, it is considerably more difficult to document the non-mechanical development and use of new methodological skills to implement these materials, not least because it is unclear how we might frame the issues that are relevant to this discussion.

What sort of evidence of methodological innovation might we look for and/or expect to find in TBLT? A plausible answer to this initial question is that the use of pair and small group work is, in itself, potentially a methodological innovation.[4] Out of this innovation may come a change in participants' classroom roles. Typically, this is thought to mean that teachers become facilitators and managers of classroom activities, not transmitters of knowledge. And students become active participants in their own learning.

How might we document methodologically-induced changes in teacher-student relationships, in particular, changes in the way power is distributed among participants? We have at least two options. We can gather empirical evidence based on transcripts of small group interaction, analyze the structural organization of the talk that occurs in small groups and compare this with the structure of the interaction that typically occurs in teacher-fronted classes. This will enable us to see if any observable changes in methodological behavior occur at certain important decision points in the discourse. Or we can ask teachers for their impressions and interpretations of what happens in their classes. Let me begin with the second option.

Teachers' Impressions

Over the years, TAs have often stated that the way in which they teach ESL at UIUC is quite different from their previous methodological practice.[5] A recurring theme in TAs' reactions has been that the heavy use of small group work that is a feature of TBLT (see, for example, Activities C and D in Section 2 of the Iceman unit) is sufficient to redistribute power in the classroom. But to what extent does small group work in which the teacher participates necessarily result in observable changes in the relationships among different participants in classroom talk?

The Organization of Teacher-Fronted and Small Group Talk

There is plenty of evidence in the literature that the student-student talk that occurs in small groups is functionally and organizationally different from that which occurs in teacher-fronted classroom talk (Long et al., 1976; Cazden, 1988). Since classroom observations of TAs over the years suggests that they do indeed promote such talk by using small group work, there is empirical support for the TAs' perception that the methodological procedures they use help create non-traditional teacher-student relationships in their classes.

But let us dig deeper. What happens when teachers are invited by learners to join a group in order to answer a question from a group member? The answer to this question is interesting because a teacher's observable behavior at this particular decision point can provide us with valuable insights into small group work: Specifically, to what extent do TAs' spontaneous response strategies actually restructure classroom relationships on a moment-by-moment basis?

In teacher-fronted instruction, classroom talk is organized in terms of a three-turn sequential structure: teachers prototypically ask question ([Q]) turns, learners provide answering ([A]) turns and teachers close sequences with evaluative commenting ([C]) turns; these turns provide feedback on the adequacy of the student's preceding [A] turn(s). By doing [Q] turns, teachers are in sequential position to control [C] turns. Thus, by implementing a [Q][A][C] turn-taking system, classroom participants collaboratively construct talk in such a way that the teacher is effectively in control of the pedagogical agenda for a lesson (Cazden, 1988; Kasper, 1985; McHoul, 1978; J. Nicholls, 1993; Pica, 1987; van Lier, 1988). This organization exemplifies a discourse of unequal power.

Arguably, TBLT should aim to break down such patterns of teacher control during at least some of the talk that occurs in the classroom. This would change traditional teacher-student relationships in the classroom - and perhaps promote greater student responsibility for their own learning - which many communicative methodologists regard as desirable. In addition, changing the practices of classroom interaction so that they more closely resemble the open-ended rules of talk found in ordinary conversation would also be valuable from an SLA point of view, since the ability to perform a broad range of communicative acts in the target language is thought to promote second language development (Pica, 1987).

Tasks that involve student-student talk in small groups go a long way toward simulating the practices of ordinary conversation: turns are not pre-allocated and all participants have co-equal rights to contribute to the conversation. We could therefore argue that a teacher who wishes to change teacher-student relationships (at least at this decision point in small group work) might promote talk that approximates the turn-taking conventions of ordinary conversation. More specifically, they might wish to respond directly to students' questions, on the grounds that a question in ordinary conversation naturally sets up the expectation that the preferred response will be an answer. This strategy would allow learners to direct both the content and sequential development of classroom talk and enable them to set their own learning agendas.

Such a sequence emerges in Excerpt 1, which comes from one of several classes that were recorded and transcribed in 1990 (see Markee, 1994 for details). The glosses in the margin of this excerpt identify the function(s) of each turn in a sequence: [Q1] = the first question in a sequence; [A]-[Q1] = the answer to question 1; and [C]-[Q1] = the commenting turn that closes the sequence initiated by question 1.

Excerpt 1
(Simplified)

1 L6:	[Q1]	what [ə]spur means? how do you pronounce it
2		s-p-u-r
3 T:	[A]-[Q1]	spu:r
4 L6:	[C]-[Q1]	spur=
5 T:	[C]-[Q1]	=//uh huh, <h>//
...		
20 L6:	[Q3]	does it have another meaning too
21 T:	[A]-[Q3]	yeah you know uh on a ho:rse (+) uhm (+) when you're
22		riding (+) you have on you::r (hh) (+) on your shoe a
23		sp//ur// and you use that to:=
24 L6:		//yeah//
25 L6:	[C]-[Q3]	=ok //I understand//
26		...

(Class 1, Group 1)

This excerpt shows that by initiating the [Q] turns on the word "spur" at lines 1 and 20, L6 is also in sequential position to close these two sequences with [C] turns at lines 4 and 25. In addition, L6 is able to set his own learning agenda throughout this piece of definition talk (i.e., finding out what "spur" means and how it is pronounced). This example of definition talk actually comprises a total of five contiguous sequences, all initiated and closed down by L6. Thus, in conversation analytic turns, Excerpt 1 clearly shows how the relationships that normally govern teacher-fronted classroom talk have been reversed at this decision point.

However, at such decision points, TAs hardly ever allow the subsequent talk to be constructed in this way; indeed, the talk illustrated in Excerpt 1 is actually the only example I have found in data from four classes where such decision points occur. TAs typically answer student questions with "counter-question" ([CQ]) turns (Nicholls, 1993). This phenomenon, which allows the teacher to re-assert control over the sequen-

tial development of the talk, is illustrated in Excerpts 2 and 3; note that Excerpt 2 involves the same teacher as in Excerpt 1, while Excerpt 3 comes from a different class. The same glosses used to analyze Excerpt 1 are used here, with the following additions: [CQ1] = the first counter-question in a sequence; [A]-[CQ1] = the answer to counter-question 1; [C]-[CQ1] = the commenting turn that closes the sequence initiated by counter-question 1; [C]/[CQ] = a turn that functions both as a commenting turn and as a counter-question turn.

Excerpt 2

1	L13:	[Q1]	... what's that mean (1) coastal vulnerability
2		-[Q1]	(1)
3	L14:	[A]-[Q1]	[f]ulnerability is:
4	L13:	[Q2]	coastal vulnera- vulnerability
5	T:	[CQ1]	what d'you think it means
6		[A]-[CQ1]	(1.3)
7	L14:	[A]-[CQ1]	uh?
8	T:	[CQ2]	what what d'you think a- where are areas of coastal vulnerability
9			(++) <h> if you think about uh:m
10	L14:	[A]-[CQ2]	it's not safe[t] (+) areas which are not safe[t] (1) right?
11	L13:	[A]-[CQ2]	it's very easy to be:: (+) damage
12	T:	[C]-[CQ2]	yea:h (+) especially by (+) water, (+) by flooding,
13			...

(Class 1, Group 4)

At lines 1 and 4 of Excerpt 2, L13 makes two requests for help with the phrase "coastal vulnerability;" instead of answering with an [A] turn (as she does in Excerpt 1), the teacher responds at line 5 by doing a first [CQ] turn. When this [CQ] turn is unsuccessful, she does a second [CQ] turn at line 8. These two [CQ] turns put L14 and L13 in the sequential position of having to do [A] turns, which they produce at lines 10 and 11 respectively; these [CQ] turns also put the teacher back in sequential position to close the sequence with a [C] turn, which she does at line 12. A similar pattern by a different teacher is also observable in Excerpt 3:

Excerpt 3

1	L6: [Q1]	there is a problem here she //doesn't//
2	L15:	//(huh h)//
3	L6:	underst(h)and
4	L7:	(huh)
5	L6: [Q1]	and we don't understand what <h>
6		//what means exactly this//
7	L15:	//why we can+t get Aus[witʃ]// (+) oh
8	L6: [Q1]	we cannot get by Ausch[v]itz
9	T: [CQ1]	ok (+) what d'you think it might mean
10	L15:	(uh huh) (+) (uh huh //h)//
11	L6: [A]-[CQ1]	//it// might [b]ean (+) probably
12		u::h we::: (+) cannot have another Ausch[v]itz again if
13		uh Germany unites o:r maybe <hh>
14	T: [C]/[CQ2]	does it mean that?
15		...

(Class 2, Phase 2, Group 2)

At lines 1, 3 and 5-6 of Excerpt 3, L6 asks the teacher to help another learner understand the phrase "we cannot get by Auschwitz." At line 9, the teacher responds to this request with a first [CQ] turn, which returns L6's question back to the learners. At lines 11-13, L6 responds to the teacher's[CQ] turn with an [A] turn of his own. However, at line 14, the teacher immediately does a second [CQ] turn; this initiates another sequence not reproduced here. However, note that the sequential position of the teacher' second [CQ] turn at line 14 immediately after L6's answer at lines 11-13 indicates that this answer is unsatisfactory. Thus, the teacher's turn at line 14 simultaneously functions as a [C] turn that closes the first [CQ] sequence initiated at line 9 and also as a [CQ] turn that initiates a second sequence whose trajectory and content are also controlled by the teacher.

In both these examples, the teachers have used a [CQ] turn-taking strategy which, whether they intended it or not, has the technical effect of re-instating a discourse of unequal power at this particular decision point in the interaction. That is, the teachers (probably quite unconsciously) take back from the students the right to control the development and content of these sequences. They thereby effectively control how the moment-by-moment[6] learning agenda is implemented. Ironically, the resulting talk is therefore empirically almost indistinguishable from the kind of talk that occurs

in "traditional" teacher-fronted classrooms, although that was almost certainly not the teachers' intention when they did these [CQ] turns.

As suggested by Fullan, therefore, the development of new materials is not necessarily complemented by the development across the board of new methodological skills. Small group work in which the teacher participates does not necessarily promote observable, methodologically-induced changes in relationships among different participants, even though such changes might be predicted from the types of tasks and techniques prescribed by the level materials. In short, this analysis suggests how very difficult it is for TAs to develop and use new methodological skills at such real time decision points in their teaching.

Dimension 3

The Development of New Pedagogical Values

If it is difficult to demonstrate the use of new methodological skills, it is even more difficult to demonstrate the development of new value systems and pedagogical beliefs. As Fullan notes, the values of what constitutes "good" teaching are quite difficult to get at (let alone change) because they are often not clearly articulated and/or are so general as to be open to a host of contradictory interpretations. At best, one can show what the process of values clarification entails.

When I asked the teacher who taught Class 1 to comment on the data contained in Excerpts 1 and 2, she replied that she was quite surprised that she had used the [A] strategy exemplified in Excerpt 1 at all. She ascribed this occurrence to being very conscious of being observed and recorded. Furthermore, she preferred the [CQ] strategy exemplified in Excerpt 2 because she felt that it helped reduce teacher talking time. She also felt that this strategy gave knowledgeable students an opportunity to practice giving definitions and that it was a student-centered way of managing the interaction—that is, it emphasized the value of cooperative rather than competitive learning. In addition, she believed that students were more likely to remember definitions they themselves produced than definitions produced by a teacher who used an [A] strategy. And finally, she noted that previous teacher training had advocated that teachers redirect questions to other students rather than answer them themselves. However, she recognized that her teacher trainers had assumed that teacher-fronted instruction was the target norm for teacher trainees.

Are these beliefs about teaching consistent with TBLT? The TA in question (who holds a Master's degree in ESL and is an experienced and very competent instructor) clearly thought so. But notice that the aim of reducing teacher talking time (TTT) is a "tenet of good teaching" that audio-lingual teachers would have had no trouble in accepting. Arguably, this tenet is an ideological left-over from previous theories about language teaching and has little to do with TBLT. Indeed, I would argue that it is quite defensible to suggest that, at decision points such as those analyzed in the previous section, teachers who wish to change the way power is distributed in a classroom should provide a direct answer to a student's question, even if this

results in a momentary increase in TTT. Notice, however, that by organizing instruction in terms of small group work activities in which she is an infrequent participant, the teacher has already succeeded in significantly reducing her own TTT in any one small group; the "problem" of too much teacher talk should therefore not be a significant concern.

Furthermore, an analysis of class transcripts shows that learners typically ask for the teacher's help with vocabulary items only after they have failed to get a satisfactory answer from their small group colleagues. This is certainly the case with L11 who, in talk that precedes Excerpt 4, has attempted to understand what the word "habitat" means three times before she asks the teacher for help.

Excerpt 4

1 L11: ok (+) excuse me (+) uh: what what does it mean hab- (+)

2 habi-

3 ...

4 T: <h> yeah what would be another word for a habitat then

5 (+) it's like (1) //it's hli-//

6 L11: //I ha//ve no idea

(Class 1, Group 4)

Thus, the teacher's use of a [CQ] strategy at line 4 of Excerpt 4 turns out to be an inefficient use of class time. As the transcript also shows, this strategy is bound to fail, since neither L11 nor anybody else in the group knows what "habitat" means. Finally, it is evident from L11's unusually frank response at line 6 that this learner merely feels frustrated when the teacher responds to her question with a [CQ] turn.

When I discussed the teacher's reactions to the conversation analysis of Excerpts 1-3 with current TAs, there was considerable disagreement as to which of the competing analyses offered (mine or the TA's) was more convincing. One TA reacted by saying that, far from avoiding the use of a [CQ] strategy, she wanted to learn how to do [CQ] turns more effectively. Her reasoning was that she wanted to improve the way she asked questions in class and that the use of this strategy was indeed an effective way of promoting collaborative learning. Others took note of the empirical data and the conversation analysis I provided but were not convinced that the use of a [CQ] strategy at such a decision point was unjustified. TAs in this group pointed out that they have to juggle with multiple contingencies in classroom talk. For them, an SLA/conversation analytic perspective such as the one I proposed fails to take into account the thousand and one practical concerns of teachers, which include getting through the material in a timely fashion, making sure that all learners get a chance to participate and other such matters. Finally, the remaining group of TAs agreed with my analysis in principle but wondered how to apply this information in their classroom practice. Furthermore, as

Excerpt 5 (an extract from a journal entry written by one of the TAs in this latter group) demonstrates, some TAs were troubled by the larger philosophical implications of this analysis:

Excerpt 5

> The goal of student-centeredness is to give students more control over their educational experiences. But there always seems to be a point at which the teacher is making the classroom do what he or she (as the teacher) perceives to be best. In other words, the teacher makes the decisions about what is best for the students. Now, I think there is definitely room for teaching professionals to be able to claim that they know what is best for their students, but I wonder about calling a classroom student-centered when the teacher is still making all the decisions that count. For example, in my classroom, my students are involved in a lot of group work, which I mandate, and which, on the whole, they have been pretty willing to do. In [their responses to questions about their learning style preferences], however, although my students said they would not like lecture all the time, they indicated that they wished that I would talk to them more, and more often.
> I'm afraid I am not making sense. It seems as though a distinction needs to be clearly made in how we are defining the expression *student-centered*. It can either mean making the student do the work of the classroom or letting student desires and goals dictate what happens in the classroom. It seems like the CQ strategy issue deals with this distinction: do we as teachers just give the students what they want (a direct answer) or do we make them work for an answer by first drawing on their own resources? I know your research shows that they have already drawn on their own resources before they ask the questions, which is a key point; nevertheless, I think there is a sense where CQing is a relatively minor manifestation of a larger conflict. Do we really want student-centeredness that gives students the power to ask us to lecture? Also, my experience as student and teacher makes me believe that people get very uncomfortable when their expectations about relationships are violated. Perhaps students, if given what they really want, will say, "You be in charge. Tell me what to do and I will do it."

This extract illustrates how difficult it is to bring about changes in our own or others' pedagogical values, *even when we are ready to change our beliefs*. All innovation ultimately depends on participants being able to engage in a meaningful process of values clarification (Henrichsen, 1989; White, 1988). In this regard, the difficulty that some TAs had in accepting the implications of the analysis of Excerpts 2-4 was caused by the fact that this analysis leads to some initially counter-intuitive methodological conclusions. More specifically, as the journal extract quoted in Excerpt 5 makes clear,

commonly used terms like "student-centered teaching" suddenly become controversial and surprisingly difficult to define.

Attempts to define key terms is a part of the process of values clarification. But this process raises some interesting problems that go to the very heart of any attempt to innovate. For example, if we cannot agree on the meaning of "everyday" terms like "student-centered teaching," it will be extremely difficult for us to change our ideas in the first place. Furthermore, it is clear that suggestions for apparently small changes (for example, using an [A] strategy instead of a [CQ] strategy) can lead to a Pandora's box of unexpected issues and problems being opened up, to which there initially seem to be no obvious solutions. Frustrating though this may be, it is nonetheless important that such issues be confronted. It is through such a process of values clarification that a complete and coherent process of curricular innovation - that is, a process of curricular innovation which encompasses all three dimensions of the innovation triangle - will emerge in the UIUC project.

Conclusion

The "innovation triangle" outlined in this article constitutes a pragmatic, yet theoretically informed, framework which outlines some of the major issues and problems that are involved in implementing curricular innovation. The use of this framework has shown that the UIUC project has only been partially successful in installing the innovation of TBLT. While we have undoubtedly been quite successful in promoting the development and use of new materials, it is less clear to what extent we have been successful in our efforts to promote the use of methodological skills that change traditional relationships between teachers and students at certain decision points in classroom talk. This is due in large part to the fact that it is difficult to reach any agreement on what kind of methodological behavior might fulfill such a goal. Finally, assessing the extent to which changes in teachers' value systems have occurred is even more problematic: It is unclear whether we should pay more attention to what people say they believe or to what they actually do.

Nevertheless language teaching professionals interested in implementing curricular innovation should not despair of ever promoting meaningful changes in their own or others' pedagogy. Innovation emerges out of attempts to resolve conflict (Candlin, 1984). Such attempts at values clarification are invariably messy; thus, the conflict to which the author of Excerpt 5 refers, for example, is an important precursor of potential change. Moreover, as this paper has demonstrated, curricular innovation *does* occur, even if it is uneven and incomplete.

The uneven development of change along the three dimensions of the innovation triangle that has been documented in this paper is quite typical of all efforts to innovate. Nevertheless, it is potentially instructive. As one of the leading authorities on promoting educational change remarks:

> While there are no resounding successes, and in many ways there is no such thing, given the perennial complexity of change, we can obtain glimpses of a more powerful future by deriving lessons from examples of partial success currently under way. It is my contention that these examples point to the central role of moral purpose [= what (language) teaching professionals seek to achieve in and through pedagogy] and change agentry, and their roles in the learning organization of tomorrow (Fullan, 1993, 59-60).

As the present and successor generations of language teaching professionals prepare for the challenges of the twenty-first century, I would like to suggest that we need to re-invent ourselves as managers of educational change. In other words, we need to devote more explicit attention to the issues and problems discussed in this paper. Specifically, we need to:

- *accept that constructing all three dimensions of the innovation triangle into a coherent program of change is an inherently messy, unpredictable and time-consuming business*

As shown by the empirical data discussed in this paper, promoting the development and use of new methodological skills and values by teachers are particularly difficult types of change to implement. Accepting this insight will mean that we have to develop a better understanding of the role of planning in (language) education while building a great deal of flexibility into our planning efforts.

- *develop an inter-disciplinary perspective on how to manage curricular innovation*

This entails borrowing and adapting ideas from disciplines such as education, language planning, sociology and development planning, all of which have more mature traditions of investigating how new ideas and practices diffuse (see Adams & Chen, 1981; Appelbaum, 1970; Baldridge & Deal, 1983; Brickell, 1969; Chin & Benne, 1976; Cooper, 1982, 1989; Cracknell & Rednall, 1986; Dow et al., 1984; Ellison & Davies, 1990; Everrard & Morris, 1990; Fullan, 1982a, 1982b, 1993; Fullan & Pomfret, 1977; Goulet, 1971; Gross et al., 1971; Havelock, 1971, 1973; Huberman, 1973; Jéquier, 1976; Kelly, 1980; Levine, 1980; Loucks et al., 1975; Miles, 1964; Milstein, 1982; Mulford, 1982; A. Nicholls, 1983; Rogers, 1983; Rogers & Shoemaker, 1971; Rondinelli, 1983; 1987; Rondinelli et al., 1990; Rudduck, 1991; Simon, 1982; Stenhouse, 1975; Zaltman & Duncan, 1977).

- *undertake action research to further understand the theoretical complexity of innovation*

Most likely, the results of this action research will highlight the fact that there are often significant discrepancies between what is intended and what is actually done. In addition, it will be important to gain practical managerial experience to learn how to make change happen.

By participating in this kind of research program, we may hope both to obtain

the glimpses of a more powerful future of which Fullan speaks and collectively to improve the quality of curriculum and teacher development in second language education. In so doing, we may even succeed in significantly improving the quality of instruction for learners. This is surely the ultimate goal of all our efforts.

Notes

1. As already noted, TBLT is new to many TAs, who typically fall in one of three categories:

> (a) those who have no knowledge of TBLT and have never used it before coming to UIUC;
>
> (b) those who sometimes use methodological procedures frequently associated with TBLT (for example, information gap or jigsaw tasks) but who do not know what the theoretical rationale for the use of such procedures might be; and
>
> (c) those who both already use and know about TBLT when they arrive on campus. Most TAs over the six years that this project has been running typically fall in either category 1 or 2.

2. The "Iceman" in the title refers to the mummified body of a Metal Age man that was found high up in the Italian Alps in 1991. The archeological importance of this find has been compared to the discovery of Tutankhamun's tomb in Egypt by Howard Carter in the 1920s.
3. This does not mean that the traditional language teaching concern with formal accuracy in writing is ignored. Problems with grammar and lexis are dealt with, but as a side product of learners' attempts to produce a coherent first draft. Thus, "grammar teaching" typically occurs as feedback provided by the teacher on the quality of individual learners' first drafts.
4. In point of fact, I am not claiming that small group work per se is an innovation for new TAs at UIUC; however, what is new for some TAs is the extent to which this procedure is used to carry instructional content in TBLT.
5. Obviously, this feeling varies with each individual, depending on how familiar or unfamiliar they actually were with TBLT before they joined the ESL courses.
6. As an anonymous reviewer pointed out, it is important to note that the teachers are responding intuitively to the situations; thus, even experienced teachers who favor the innovation may behave in the same way.

References

Anderson, A. & Lynch, T. (1988). *Listening.* Oxford: Oxford University Press.

Appelbaum, R. P. (1970). *Theories of social change.* Chicago: Markham.

Baldridge, J. V. & Deal, T. E. (1983). The basics of change in educational organizations. In J. V. Baldridge & T. E. Deal (Eds.), *The dynamics of organizational change in education* (pp. 1-11). Berkeley, CA.: McCutchan.

Breen, M. P. (1984). Process syllabuses for the language classroom. In C.J.Brumfit (Ed.), *General English syllabuses design (ELT Documents 118, pp. 47-60).* Oxford: Pergamon and the British Council.

Breen, M. P. (1987). Learner contributions to task design. In C. N. Candlin & D. F. Murphy (Eds.), *Language learning tasks, Lancaster practical papers in English language education, 7,* 23-46.

Breen, M. P. & Candlin, C. N. (1980). The essentials of a communicative curriculum. *Applied Linguistics, 1*(2), 89-111.

Brickell. (1969). Appraising the effects of innovation in local schools. In R.W. Tyler (Ed.), *Educational evaluation: New roles, new means (pp. 284-304).* Chicago, IL: National Society for the Study of Education.

Brindley, G. (1984). *Needs analysis and objective-setting in the Adult Migrant Education Program.* Sydney: Adult Migrant Education Service.

Brock, C.A. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly, 20,* 47-59.

Broersma, D. (1993a). *The iceman speaks.* Unpublished task-based unit contributed to the UIUC bank of materials, University of Illinois at Urbana-Champaign.

Brown, G. & Yule, G. (1983). *Teaching the spoken language.* Cambridge: Cambridge University Press.

Candlin, C. N. (1984). Syllabus design as a critical process. In C. J. Brumfit (Ed.), *General English syllabus design* (ELT Documents *118*, pp. 29-46). Oxford: Pergamon and the British Council.

Candlin, C. N. (1987) Towards task-based language learning. In C. N. Candlin & D. F. Murphy (Eds.), *Language learning tasks, Lancaster practical papers in English language education,* 7, 5-22.

Cazden, C. (1988). *Classroom disclosure: The language of teaching and learning.* Portsmouth, NH: Heinemann.

Chin, R. & Benne, K. D. (1976). General strategies for effecting changes in human systems. In W. G. Bennis, K. D. Benne, R. Chin & K. E. Corey (Eds.), *The planning of change* (3rd ed., pp. 22-45). New York: Holt, Rinehart and Winston.

Clarke, D. F. (1991). The negotiated syllabus: What is it and how is it likely to work? *Applied Linguistics, 12*(1), 13-28.

Cooper, R. L. (1982). A framework for the study of language spread. In R. L. Cooper (Ed.), *Language spread: Studies in diffusion and social change* (pp. 5-36). Bloomington, IN: Indiana University Press and Center for Applied Linguistics, Washington, DC.

Cooper, R. L. (1989). *Language planning and social change.* Cambridge: Cambridge University Press.

Cracknell, B. E. & Rednall, J. E. (1986). *Defining objectives and measuring performance in aid projects and programmes.* London: Overseas Development Administration.

Doughty, C. & Pica, T. (1986). Information gap tasks: Do they facilitate second language acquisition? *TESOL Quarterly, 20*, 305-325.

Dow, I. I., Whitehead, R. V. & Wright, R. L. (1984). *Curriculum implementation: A framework for action.* (ED 256715), Alexandria, VA: ERIC Document Reproduction Service.

Ellison, L. & Davies, B. (1990). Planning in education management. In B. Davies, L. Ellison, A. Osborne & J.West-Burnhan (Eds.), *Education management for the 1990s* (pp. 31-44). London: Longman.

Everrard, B. E. & Morris, G. (1990). *Effective school management* (2nd ed.). London: Paul Chapman.

Fullan, M. (1982a). *The meaning of educational change.* New York: Teacher's College Press, Columbia University.

Fullan, M. (1982b). Research into educational innovation. In H. L. Gray (Ed.), *The management of educational institutions* (pp. 245-262). Lewes: The Falmer Press.

Fullan, M. (1993). *Change forces: Probing the depths of educational reform.* London: The Falmer Press.

Fullan, M. & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research, 47*(2), 335-393.

Goulet, D. (1971). *The cruel choice.* New York: Academic Press.

Gross, N., Giacquinta, J. B. & Bernstein, M. (1971). *Implementing organizational innovations.* New York: Basic Books.

Havelock, R. G. (1971). The utilization of educational research and development. *British Journal of Educational Technology, 2*(2), 84-97.

Havelock, R. G. (1973). *The change agent's guide to innovation in education.* Englewood Cliffs, NJ: Educational Technology Publications.

Henrichsen, L. E. (1989). *Diffusion of innovations in English language teaching: The ELEC effort in Japan, 1956-1968.* New York: Greenwood Press.

Huberman, A. M. (1973). *Understanding change in education: An introduction.* Paris: OECD.

Jéquier, N. (1976). The major policy issues. In N. Jéquier (Ed.), *Appropriate technology: Problems and promises* (pp. 16-112). Paris: OECD.

Johnson, K. (1982). Five principles in a "communicative" exercise type. In K. Johnson (Ed.), *Communicative syllabus design and methodology* (pp. 163-175). Oxford: Pergamon.

Kasper, G. (1985). Repair in foreign language teaching. *Studies in Second Language Acquisition, 7*(2), 200-215.

Kelly, P. (1980). From innovation to adaptability: The changing perspective of curriculum development. In M. Galton (Ed.), *Curriculum change: The lessons of a*

decade (pp. 65-80). Leicester: Leicester University Press.

Kennedy, C. (1988). Evaluation of the management of change in ELT projects. *Applied Linguistics, 9*(4), 329-342.

Kiniry, M. & Strenski, E. (1985). Sequencing expository writing: A recursive approach. *College Composition and Communication, 36*(2), 191-202.

Krashen, S. D. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon.

Krashen, S. D. & Terrell, T. D. (1983). *The natural approach.* Hayward, CA: The Alemany Press.

Kumaravadivelu, B. (1991). Language learning tasks: Teacher intention and learner interpretation. *ELT Journal, 45,* 98-111.

Kumaravadivelu, B. (1993). The name of the task and the task of naming: Methodological aspects of task-based pedagogy. In G. Crookes and S. Gass (Eds.), *Tasks in pedagogical context: Integrating theory and practice (pp. 69-96).* Philadelphia: Multilingual Matters.

Levine, A. (1980). *Why innovation fails.* Albany, NY: State University of New York Press.

Long, M. H. (1981). *Input, interaction and second language acquisition.* Unpublished Ph.D. dissertation, University of California at Los Angeles.

Long, M. H. (1985). A role for instruction in second language acquisition. In K. Hyltenstam & Pienemann, M. (Eds.), *Modeling and assesing second language acquisition* (pp. 77-100). Clevedon, Avon: Multilingual Matters.

Long, M. H. (1989). Task, group, and task-group interactions. *University of Hawaii Working papers in ESL, 8*(2), 1-26.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Boot, R. B. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-51). Amsterdam and Philadelphia: John Benjamins.

Long, M. H., Adams, L., McLean, M. & Castanos, F. (1976). Doing things with words - Verbal interaction in lockstep and small group classroom situations. In J. F. Fanselow & R. Crymes (Eds.), *On TESOL '76 (pp. 137-153).* Washington, DC: TESOL.

Long, M. H. & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26* (1), 27-56.

Long, M. H. & Crookes, G. (1993). Units of analysis in syllabus design: The case for task. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9-54). Clevedon, Avon: Multilingual Matters.

Loschky, L. & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes and S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice (pp. 123-167).* Clevedon, Avon: Multilingual Matters.

Loucks, S. F., Newlove, B. W. & Hall, G. E. (1975). *Measuring levels of use: A manual for trainers, interviewers, and raters.* Austin, TX: The Research and Development Center for Teacher Education, University of Texas at Austin.

Markee, N. P. P. (1993). The diffusion of innovation in language teaching. *Annual Review of Applied Linguistics, 13*, 229-243.

Markee, N. P. P. (1994). Toward an ethnomethodological respecification of second language acquisition studies. In A. Cohen, S. Gass, & E. Tarone (Eds.), *Research methodology in second language acquisition* (pp. 89-118). Hillsdale, NJ: Lawrence Erlbaum.

McHoul, A. (1978). The organization of turns at formal talk in the classroom. *Language in Society, 19*, 349-377.

Miles, M. B. (1964). Educational innovation: The nature of the problem. In M. B. Miles (Ed.), *Innovation in education* (pp. 1-48). New York: Teacher's College Press.

Milstein, M. M. (1982). Training internal change agents for schools. In H. L. Gray (Ed.), *The management of educational institutions* (pp. 163-178). Lewes, Sussex: The Falmer Press.

Mulford, W. (1982). Consulting with educational systems is about the facilitation of coordinated effort. In H. L. Gray (Ed.), *The management of educational institutions* (pp. 179-218). Lewes, Sussex: The Falmer Press.

Nicholls, A. (1983). *Managing educational innovations.* London: Allen and Unwin.

Nicholls, J. (1993). Exchange structure in the ESL classroom: Q-A-C and Q-CQ-A-C sequences in small group interaction. *Pragmatics and Language Learning Monograph Series, 4*, 183-193

Nunan, D. (1987). Communicative language teaching: Making it work. *English Language Teaching Journal, 41*, 136-185.

Nunan, D. (1988). *Syllabus design.* Oxford: Oxford University Press.

Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge: Cambridge University Press.

Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL Quarterly, 25*(2), 79-296.

Nunan, D. (1993). Task-based syllabus design: Selecting, grading and sequencing tasks. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9-34). Clevedon, Avon: Multilingual Matters.

Pica, T. (1987). Second language acquisition, social interaction and the classroom. *Applied Linguistics, 8*, 3-21.

Pica, T., Kanagy, R. & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice (pp. 9-34).* Clevedon, Avon: Multilingual Matters.

Prabhu, N. S. (1987). *Second language pedagogy.* Oxford: Oxford University Press.

Rogers, E. M. (1983). *The diffusion of innovations* (3rd ed.). London: MacMillan and Free Press.

Rogers, E. M. & Shoemaker, F. (1971). *Communication of innovations: A cross-cultural approach* (2nd ed.) New York: Free Press.

Rondinelli, D. (1983). *Development projects as policy experiments.* London: Methuen.

Rondinelli, D. (1987). *Development administration and US foreign policy.* Boulder: Lynne Rienner

Rondinelli, D., Middleton & Verspoor (1990). Planning educational reforms in developing countries. Durham: Duke University Press.

Rudduck, J. (1991). *Innovation and change.* Milton Keynes: Open University Press.

Simon, R. I. (1982). Mysticism, management and Marx. In H.L. Gray (Ed.), *The management of educational institutions* (pp. 63-85). Lewes, Sussex: The Falmer Press.

Stenhouse, L. (1975). *An introduction to curriculum research and development.* London: Heinemann.

Tetenbaum, T. J. & Mulkeen, T. A. (1986). Designing teacher education for the twenty-first century. *Journal of Higher Education, 57*(6), 621-636.

van Lier, L. (1988). *The classroom and the language learner.* London: Longman.

White, R. V. (1988). *The ELT curriculum: Design, innovation and management.* Oxford: Basil Blackwell.

Wilkins, D. A. (1976). *Notional syllabuses.* Oxford: Oxford University Press.

Zaltman, G. & Duncan, R. (1977). *Strategies for planned change.* New York: John Wiley and Sons.

Author

NUMA MARKEE, Assistant Professor, Division of English as an International Language, University of Illinois at Urbana-Champaign, 3070 Foreign Languages Bldg, Urbana, IL 61801. Specializations: management of curricular innovation, SLA/classroom research, language planning.

Applied Language Learning
1994, Vol. 5, No. 2, pp. 31-56

Computer-Assisted Language Instruction in Cooperative Learning

Lydie E. Meunier
The University of Tulsa

This article addresses the renewed debate on cooperative learning in foreign language classrooms. First, it reports on the rationale for communicative language teaching in foreign language (FL) instruction. The focus then shifts to research on potential drawbacks of communicative language teaching: i.e., the selection of task-types, the lack of sociolinguistic authenticity, and the low quality of language input during communicative activities. The second part of the article focuses on the role computers can play in cooperative learning today, and how they can offset the drawbacks observed in "conventional" communicative activities.

Introduction

During the previous decade, the principles of communicative language teaching paved the way to a greater use of techniques involving group activities in FL classrooms. Yet, the concept of communicative language teaching is still being somewhat misunderstood. Group activities are not simply a matter of putting students together during grammar-based or structural activities, opportunities must also be given to facilitate open communication and interaction in the target language (Brown, 1991). Because of the non-threatening atmosphere inherent in the cooperative context, the quality of the communication among FL students has been shown to be better than in teacher-centered classrooms (Seliger, 1983). More recent studies suggest, however, that communicative language teaching lack sociolinguistic authenticity: While speaking with peers, students are provided with poor quality input, also referred to as "interlanguage input," which is likely to lead to fossilization (Aston, 1986; Ellis, 1992a, 1992b; Lightbown, 1992). Yet, while fossilization has been found to be a possible shortcoming of communicative language teaching; communicative activities in the FL classroom are still viewed as a significant predictor which fosters motivation, attitude, and confidence; essential factors for FL acquisition (Aston, 1986; Ellis, 1992a, 1992b; Lightbown, 1992). Thus educators who promote active learning and teach profi-

© 1994, Lydie Meunier

707

ciency will surely reject a suggestion by Wong-Fillmore (1992) for a teacher-centered approach to ensure quality input. Rather, the question that needs to be addressed is: How can we improve communicative language teaching by providing quality input? In an attempt to answer this question, teachers have used more and more communicative activities within the framework of computer-assisted language instruction (CAI), and more especially the use of computer games, electronic mail, and live-action simulations brought to class through motion videos. Indeed, computers can now talk and provide students with high quality language input. The integration of computer-based group activities into lesson plans also implies that the computer lab should be expanded to classrooms, and that some classrooms should be equipped with "desk-built-in-platforms." Keeping in mind the importance of integrating a higher quality input during group activities, as well as the relevance of cooperative learning for affective, social, and motivational reasons this article will first review the rationale for cooperative learning before explaining in detail the new role that computer-assisted language learning (CALL) can play in communicative language teaching.

The Rationale for Communicative Language Teaching

Positivism versus Constructionism

Cooperative learning in FLs has roots in studies in educational psychology. Traditionally, education has been based on a positivist philosophy whereby scientific methods are applied in a linear way: There is transfer of knowledge from the scientist (teacher) to the novice learner. Knowledge is acquired through a system of extrinsic rewards (Freire, 1970; Reiser, 1987; Winn, 1989; Kubota, 1991). Within these pedagogical principles, taught knowledge is a "truth" that students must memorize. Questioning the "truth" is perceived as a challenge to the instructor's authority: Discussions are barely acceptable. In a constructionist approach, however, knowledge is "constructed" by every student with the help of the teacher and by means of various types of interaction. According to the constructivists, knowledge has no absolute validity (Goodman, 1984; Kubota, 1991), and learning is the result of an underlying cognitive development which, according to Piaget (1977) undergoes various phases: (1) assimilation (i.e., the raising of hypotheses to understand underlying systems); (2) disequilibrium (i.e., the necessary condition to readjust some of the hypotheses); (3) adaptation, a phase that eventually leads to standard norms. Likewise, Vygotsky (1978) states that human beings do not learn only by an independent ability, they also learn through an ability acquired in social interaction, a learning approach referred to as the "zone of proximal development."

Toward a "Productive Conviviability"

The relevance of Vygotskyan theory for the teaching of FLs has been recognized and serves as a valuable model because of its interactionist foundation:

> Following the Vygotskyan model of learning, students are regulated by three factors: the object (i.e., the forms and structures of the target language), the other (the teacher and others in the class) and the self (each learner)... By engaging the students in dialogs, the teacher helps to achieve a balance of all three regulatory forces. (Di Pietro, 1987, p. 15)

Clearly, this concept of interactive learning lends itself to linguistic exchanges, an ideal schema for FL studies. Rivers (1987) made it very clear that language without communication loses its very essence, and does not deserve to be called "language" when primarily handled like mathematics. Language has a special dimension that other academic topics lack: It helps us relate to other people in a society, and to the world at large. Yet, the teaching of FLs still too often neglects the communicative aspect of language. Early theoreticians of the proficiency movement (see Omaggio-Hadley [1993,] for more details) saw that audiolingualism had failed to teach students how to communicate effectively in the target language in real-life contexts. Slowly, the FL profession has moved from what Kramsch (1987) called "institutional productivity" to "productive conviviability" (p. 28), thus preparing students to socialize with people using the language. With the emergence of proficiency and communicative language teaching, language functions defined as purposes of language use have become the ultimate and most important objectives of FL pedagogy. Consequently, communicative language teaching does not raise the question "What are the structural details to be taught?" but rather "What is the communicative purpose of a specific text?"

Arguments in Favor of Group Activities

Advocates of the early proficiency-oriented movement (Omaggio-Hadley, 1993) have stressed three arguments in favor of group activities in the FL classroom: (1) group work helps students to overcome the anxiety of speaking a foreign language (speaking with a peer is less embarrassing than speaking to a teacher in front of the whole class); (2) group activities give students the opportunity to use the FL; and (3) the quality of the communication has been shown to be better than in teacher-centered classrooms because of the non-threatening atmosphere inherent in the communicative context. Likewise, Littlewood (1974) indicated that developing students' communicative competence paved the way to more grammatical accuracy in foreign language learning. Similar results have been reported by Long and Porter (1985) who found that group work: (1) provides students with substantially more language practice opportunities; (2) overall quality of student talk improves because students feel more confident expressing themselves. As Davies and Omberg (1987) posit, student talk encourages them to move on to more accurate language use because of a growing interest in the language they study. The authors also point out that group work indi-

vidualizes instruction; promotes a positive affective climate, and fosters motivation. These findings did not only account for the students' motives to learn a second language, but also for the importance of the socio-affective instructional environment as a significant factor for maintaining or altering the initial motivations. In a study on task motivation through communicative activities, Pica (1987) suggests that even if the students' motivation is strong when they first enroll in a foreign language class, the lack of motivating activities in the instructional environment can hinder language acquisition. More specifically, Pica mentions that classrooms fall short in providing the type of social and psychological interaction necessary for genuine exchanges because the instructional environment is still primarily perceived as a place where there is a lack of equality between instructors and students.

Task Types and Intrinsically Motivating Activities

In a more recent study, Pica, Holliday, and Lewis (1991) found that the amount of communication between partners correlated with how open a task is. Indeed, students still often experience a FL education essentially based on highly structured tasks: activities, even if done in groups, are still often designed to get specific types of information, allowing very little flexibility. An instructional approach based on the pleasure of discovering and constructing knowledge during the learning process has not yet gained much face validity. Research by Deci and Ryan (1985) indicates, however, that education is the result of both extrinsic or intrinsic motivation. When intrinsically motivated, students tend to excel at certain tasks because they are fun or interesting. Extrinsic motivation, on the other hand, only takes place when rewards such as grades, or even threats are used as incentives. Research suggests that classroom environments are unbalanced because there is still too much emphasis on extrinsic motivation:

> In teacher-oriented interaction the teacher selects the next speaker and automatically selects him - or herself for the succeeding turn. There is little motivation for students to listen to one another, and the only motivation for students to listen to the teacher is the fear of being caught short on an answer. (Kramsch, 1987, p. 22)

The prevalence of the well-known IRF (Initiation-Response-Feedback) exchange in teacher-centered classes have shown to restrict second language (L2) learners to a minimum number of speech acts (Sinclair & Coulthard, 1975) and to limit L2 production (Seliger, 1983; Pica & Doughty, 1985a). As such, group work is believed to offer more opportunities for students to foster language production.

Keeping in mind the distinction between intrinsic and extrinsic motivation, recent trends in FL education have favored student-centered interactions. Instructors have introduced more meaningful and challenging exchanges in FL classes. They have come up with activities aiming at exchanging information which is relevant to the

FL learner: "To increase learner motivation, there should be an increased emphasis on content, - i.e., on worthwhile, thought- and emotion-provoking information and interaction..." (Schulz, 1991, p. 175). Ely (1986a) indicated that motivation and situational comfort provided by communicative language activities can play a stronger role in the process of language development than language aptitude and intelligence. Along the same lines, Terrell (1977, cited in Schulz, 1991, p 172) observed that "the primary factors which influence L2 acquisition are affective and not cognitive", and "language learning takes place, when there is a real need and motivation for it." Wilga Rivers (1983) reminds FL professionals that language is a "vehicle for the communicating of ideas, emotions and experiences... [and that] we call languages "dead" when they no longer function this way" (p. 77).

Personalizing L2 Instruction

Language learning within the framework of communicative language teaching (Higgs & Clifford, 1982) is defined in terms of goals, i.e., fostering the learner's ability to communicate functionally in the target language, adding a sociolinguistic aspect to the classroom through the use of authentic materials. Communicative language teaching has been described as an eclectic and flexible approach which borrows various techniques from other methods. Techniques are adopted on the basis of the students' and the instructor's personality as well as on the basis of the material to be taught. In other words, instructors in a communicative language teaching environment are extremely sensitive to how students respond to the various techniques being used, and they aim at optimizing situational comfort in order to ensure maximum language development. Rivers (1983) describes this instructional environment in the following terms:

> Cooperative learning implies full participation of both teacher and student and the interaction of student with student. It implies participating in planning and the opportunity to make effective choices. It implies student helping student, student helping teacher; it implies small group activity, large group instruction, interacting in pairs, or leaving another individual alone if that is what he or she prefers. It implies sharing what one has discovered with others. This is surely what education, as opposed to instruction, is about: "a deliberate and conscious sharing of responsibility for learning. " (p. 78)

In other words, cooperative learning is not simply a matter of putting students together to make them speak. Communicative techniques are primarily meant to structure the classroom environment to facilitate interaction in the target language: "interaction is the fundamental fact of pedagogy, and successful pedagogy involves the successful management of classroom interaction" (Allwright, 1984, p. 156).

Management of Classroom Interaction

According to Allwright, group activities must be carefully organized to enhance maximum language production, namely:

- •objectives must be set;
- •decision-making must be encouraged;
- •negotiations of meaning must be fostered;
- •activities must be monitored;
- •students must also be aware that their performance is being evaluated (Johnson & Johnson, 1985, 1986).

Kagan (1987) mentions four distinct types of cooperative learning activities: peer tutoring, jigsaw (e.g. a student describing a picture to a partner whose task is to draw it), cooperative projects, and cooperative interactions. Kagan indicates that in the framework of communicative language teaching, students progress at their own pace through individualized participation while being motivated by the team's achievements.

Within this communicative schema, Scarcella and Oxford (1992) referred to teachers as "facilitators of group dynamics" who "view errors as a creative aspect of second language development" (Scarcella & Oxford, 1992, pp. 5, 26). Indeed, FL instructors have to be aware of the fact that they typically view language as a set of conventions, whereas foreign language learners tend to view language as creation. If creation, also referred to as "risk-taking," is not encouraged among students, a psychological "blockage" prevents them from natural L2 development. Tarone (1982) suggests that too much concern about making mistakes contributes negatively to language performance: A high concern for language form leads to more anxiety and less self-confidence; the higher the level of criticism on the part of the teacher, the lower the amount of risk-taking (Tarone, 1982; Pica & Doughty, 1985b). While an immediate perfectionism was the primary idea of traditional FL teaching, the acknowledgment of a natural cognitive development taking place during interactions became the leading principle of communicative teaching. The old positivist perspective had rejected errors as a normal part of cognitive development. Today, pedagogues teach by fostering their students' creations, whether these creations are normative or not. Thus students are encouraged to take risks, to invent and discover, rather than process and memorize knowledge, until they reach the desired norm. These principles are supported by research in FL acquisition which has shown that communication helps in constructing and reconstructing hypotheses about the language, as well as in modifying interlanguage through negotiated exchanges (Long, 1985; Schachter, 1986). Within this teaching paradigm based on communication and on a non-prescriptive language view, students cognitively benefit from the shift in their role, i.e. from passive information consumers to active knowledge builders.

Empirical Research in Cooperative Learning and Task Continuum

Empirical studies investigating cooperative learning situations have shown interesting results. Seliger (1983) found that the learning rate was faster among students who were interacting with peers than students passively receiving knowledge from a lecturer. Likewise, students involved in a cooperative learning environment showed a higher degree of achievement and a better attitude toward the language they were learning (Verano, 1987; Green, 1993). Pica and Doughty (1985a) observed that ESL students showed more negotiation of meaning in group work than in teacher-centered classrooms. This finding is important because negotiations of meaning (also defined as "talk adjustments") are considered to foster comprehensible input, a precondition to language acquisition (Krashen & Terrell, 1983). Yet, the relevance of negotiations of meaning have contributed less to language development in one-way tasks than in two-way tasks (Doughty & Pica, 1986). One-way tasks (jigsaw, e.g. a student describing a picture to a partner whose task is to draw it) involve fewer negotiations of meaning than two-way tasks (solving a problem together, e.g. role-playing). Yet, one-way tasks have shown more negotiations than teacher-fronted classes (Pica & Doughty, 1985a). Likewise, Pica (1987) finds that decision-making tasks do not promote second language acquisition as much as two-way conversational exchanges because they do not require as much negotiation of meaning. These studies show that group work provides L2 learners with significantly more individual practice opportunities and incentives than teacher-centered classes. Likewise they recommend that in group work the communicative nature of language use should be as authentic as possible. Otherwise, if a task is perceived as too artificial, students lose interest (Doughty & Pica, 1986; Pica, 1987; Verano, 1987).

Ballman (1988) examined the effectiveness of cooperative learning among university students in a first-semester Spanish course and compared them to students who were assigned individual practice sheets. Results of this study showed no significant difference between achievement of the students involved in cooperative or individual tasks, although students preferred to work in groups. It should be pointed out, however, that the group activities in this study consisted of structured mechanical exercises during which students spoke about language and did not use language in meaningful communicative interactions. One can wonder, for instance, how the results would have been different if the experiment had been conducted with role-play situations instead of mechanical exercises. Ballman's experience (1988) gives additional weight to Tarone's (1982) observation that when instruction focuses on form, there is a smaller amount of communicative interaction.

Tarone's (1982) study showed that when adult ESL learners of different language backgrounds are involved in communicative activities, they acquire grammatical morphemes in a natural sequence and that the language they use reflects a coherent language system. Yet, when the same learners were asked to perform highly structured tasks, Tarone found that they lost their linguistic coherence, became hyper correc-

tive, and lost confidence. Her study suggests that natural language use fails to occur when the focus is on manipulating forms rather than expressing meaning. Bygate (1988) and Green (1993) have confirmed Tarone's study by showing that interactions based on meaningful exchanges are more significant to the learners' needs than language interactions based on artificial tasks. Bygate and Green's studies focused on the student's perception and judgments of accuracy-based practices and communicative activities. The analysis of questionnaires indicated that students were more attracted to real-language use than to practice sessions on formal grammar.

A more recent study by Brown (1991) found that the degree of "tightness" or "openness" of tasks is not the only criterion to affect language use. Significant differences were also observed in the quality of language output between so-called "procedural" and "interpretive" tasks. In procedural tasks, discussions center on decisions to be made. In interpretive tasks, students interpret a situation according to their own understanding of the world. Interpretive tasks have been found to lead to the use of more complex language among L2 learners than procedural tasks (Brown, 1991). Brown (1991) suggests that procedural tasks often require such a rigorous reasoning that language use may be impeded, whereas interpretive tasks allow more personal involvement and less inhibition.

Pica and Doughty (1985b) mention that task types have a more significant effect than interaction types, i.e. group vs. teacher-fronted classrooms, on the quality of language input and language negotiations. They suggest that more structured activities are important to help students learn to monitor their speech. When tasks are too open, the accuracy of language output declines. Another study conducted among students learning Spanish at the college level (Hartl, 1984) demonstrated a positive effect of cooperative learning activities on both achievement and attitude.

These studies suggest that along the interaction continuum, tasks performed in an instructional environment vary on three dimensions, each of which affects the degree of interactivity. Activities can be either teacher-controlled or group-controlled, tight or open, procedural or interpretative. The results imply that task selection should be done in full awareness of these differences and of their impact on students. What, then, are the most efficient tasks for language learning? Rivers states that FL students, as they progress through their learning, need to experience a continuum ranging from mechanical tasks, to semi-mechanical, to structured and open communicative activities, and that all types of tasks have a role to play in language learning. Rivers indicates that in the process of language acquisition, there is a "skill-getting" phase accounting for the learning of forms prior to reaching a "skill-using" phase necessary for actual language acquisition (Rivers, 1983, 1987). She further implies that the nature of tasks must be adapted to the language skill at stake and that mechanical or conversational activities must be selected according to the learners' needs, while still keeping in mind that the final goal of the process is communication.

Drawbacks of Communicative Language Teaching

Need of a "Silent Period" for Some Learners

Recent cognitive theories (McLaughlin, 1987) support Rivers' distinction between skill getting and skill using. The cognitivists make a distinction between the acquisition of procedural knowledge (skill getting through mechanical and semi-mechanical activities) and the acquisition of declarative knowledge (skill using by actually speaking and experiencing situations). Within this theoretical framework, language tasks are believed to activate nodes (or cognitive units) assigned with vectorial weights calculated on the frequency of previous neuron excitements. In other words, the activation of cognitive units depends on previous sociolinguistic experiences acquired through skill-using tasks. As such, skill-using tasks predetermine an automatic "spread of activation" whereby various types of declarative and procedural knowledge are retrieved at the same time for the purpose of language production, i.e., the skill-using phase.

McKenna (1987) observed that successful ESL speakers learn how to become involved in meaningful exchanges but that they first need to go through a period of self-centeredness while struggling to learn the vocabulary and forms of the language. This period of self-centeredness coincides with the so-called "silent period" (Krashen, 1985; Saville Troike, 1988). Brown (1987) describes this phase as "the schizophrenic period of culture shock and language learning... a period during which the learner will either sink or swim" (Brown, 1987, p. 131). Krashen (1985) explains such situations by the fact that language learners go through a silent period at the beginning of their learning experience. Krashen argues that during this silent period, a skill-getting phase, language learners start internalizing the comprehensible input they receive. Once learners feel they have acquired enough confidence and enough competence, speech emerges. Saville-Troike (1988) has shown that the silent period was not necessarily a period of categorical silence, and that some L2 learners were in fact using private speech as a learning strategy before directing their speech to others. According to Bygate (1988), this period of silence and of private speech depends on the student's personality: Teacher-fronted teaching can be more comforting for introverted students who need a silent period to internalize new materials. Yet, extroverted students need to try out new materials as early as possible in communicative situations. As such, group works may not always be appropriate for every type of early language development. Likewise, it may not be adequate to all personality types. Some students, as shown by Krashen and Saville-Troike, learn better by listening than by speaking, at least at certain phases of their development.

Lack of Sociolinguistic Authenticity and of Quality Input

Besides questioning the appropriateness of group work, other studies have also criticized communicative language teaching for its lack of sociolinguistic authenticity and for the poor quality input it provides. Ellis (1992a) mentions that interlanguage input is not ideal for language acquisition because it lacks quality and may not even be always comprehensible. Likewise, Aston (1986) points to the danger of language fossilization because transferred language patterns are used across students sharing the same native language. Interlanguage patterns are reinforced to the point they may fossilize. Longitudinal studies reported by Lightbown (1992) have confirmed Aston's observation: "Incorrect hypotheses about the target language find massive confirmation in the interlanguage-based speech of classmates" (Lightbown, 1992, p. 191) thus leading to a large-scale fossilization of erroneous language forms. Ellis (1992b) conducted a longitudinal study of two ESL beginners and their language development over a period of four semesters. He observed that while the interactive context of a classroom can satisfy some communicative needs, as well as promote meaningful linguistic exchanges and substantial language development, the instructional context is inherently limited in the range of genuine sociolinguistic situations it can provide. Due to this limitation, students cannot reach "full target language norms" (Ellis, 1992b, p. 20). Ellis thus proposes to improve communicative language teaching by "creating such [sociolinguistic] needs artificially" in the classroom (Ellis, 1992b, p. 20).

Breen (1985) agrees with Ellis' observations, claiming that "the classroom may be a relatively inefficient environment for the methodical mastery of a language system, just as it is limited in providing opportunities for real world communication in a new language" (Breen, 1985, p. 154). Yet, neither Ellis nor Breen specify how they would plan to go about integrating greater sociolinguistic authenticity into the FL classroom. Although the lack of sociolinguistically authentic contexts in an instructional environment is a potential shortcoming, it should be noted that Ellis' (1992) data were collected over a period of only four semesters, a span of time which may not be long enough to expect the acquisition of a "full target language norm". Considering that students improved considerably in the way they were manipulating the language, the interactive context in which they learned English actually proved to be beneficial.

Wong-Fillmore (1992) suggests that coming back to a more teacher-centered mode is one possible solution for providing quality input when "the only speaker of the language is the teacher" (Wong-Fillmore, 1992, p. 46). Yet, it has been mentioned earlier that exchanges in teacher-fronted classes are limited by the IRF (Initiate-Response-Feedback) structure, and that even if quality input is provided by teachers, the admittedly few sociolinguistic contexts introduced in role-play situations and other group activities would entirely disappear and the range of practice opportunities would be even narrower. On the "continuum that extends from instructional to natural discourse" (Kramsch, 1985, p. 169), L2 instruction would certainly regress.

While some recent empirical data indicate that fossilization can be a serious shortcoming of communicative language teaching (Aston, 1986; Ellis, 1992a; Ellis, 1992b; Lightbown, 1992), the same studies maintain the usefulness of communication so that motivation, attitude, and confidence— two essential factors in L2 acquisition— remain high among L2 students.

Before investigating how computers can offset drawbacks noticed in FL communicative teaching, it is useful to summarize briefly the principles for cooperative learning reviewed so far. It has generally been found that:

> (1) cooperative learning fosters motivation and learning outcomes in L2 classes;
> (2) communication is the *reason d'être* of language use (a language not used for communication is called a "dead" language);
> (3) learning is not only an individual, but also a social and creative process;
> (4) while cooperative tasks have been criticized for not providing quality input and for causing fossilization, cooperative contexts have been shown to improve risk-taking, self-esteem, positive attitude and motivation, all psychological constructs known to be relevant factors in foreign language development.

CAI and Communicative Language Teaching

What is Cooperative CAI?

Early research on attributes of Computer-Assisted Instruction (CAI) (a general term that applies to all academic fields) praised the computer's ability to accommodate individual learner differences. More recent research indicates that computer-assisted instruction can enhance cooperative learning as well. Johnson and Johnson (1986) studied the limitations of computer-assisted individual learning and concluded that:

> (1) the lack of social interaction leads to boredom and frustration;
> (2) the students' attitude and motivation tend to drop when working alone on a computer;
> (3) being alone on a computer does not give students the opportunity to summarize and explain orally what they have learned (oral explanations enhance higher cognitive reasoning);
> (4) computers do not show how to develop problem-solving strategies, whereas peers do;
> (5) computers only provide feedback on what is entered (i.e., the end product), whereas peers evaluate and praise the whole process of the learning approach;

> (6) creativity is limited when working alone on a problem-solving task.

It has been observed that certain computer activities can generate group cooperation and discussion. For example, computer-based cooperative problem-solving has shown to provide motivating experiences for students (Fisher, 1984; Guerrero, Eisler, Budin, & Lane, 1990; Webb, 1984; Webb, Ender, & Lewis, 1986; Webb & Kenderski, 1984). Krendl and Lieberman (1988) showed that individual-computer assisted instruction (CAI) did not enhance learning as much as did cooperative CAI. Other studies (Meravech, Stern, & Levita, 1987; Meravech, Silber, & Fine, 1991) reach the same conclusions and indicate that cooperative CAI alleviates the anxiety of low ability students more than individual CAI. A more recent study confirms these findings, indicating that pairing at the computer facilitated procedural and content information-giving while fostering a strong positive effect (Feldmann, Fish, Friend, & Bastone, 1991). Johnson, Johnson, and Stanne (1985) compared three types of CAI: cooperative, competitive, and individual. Their findings indicate that cooperative CAI produces better results on achievement tests, a better retention of factual knowledge, and a better use of factual knowledge for problem-solving situations than in individual CAI and competitive CAI. The results between competitive and individual CAI did not show any significant difference but students in cooperative CAI manifested a more positive attitude toward computers; they were faster and more accurate in problem-solving situations, and their conversations were more focused.

Carrier and Sales (1987) conducted studies among students working with computers either individually or in pairs. Although no difference was found in immediate and delayed posttest between individual and cooperative CAI (Carrier & Sales, 1987; Guntermann & Tovar, 1987), the study of verbal interactions in groups showed that students were on task 80 percent of the time and that paired students exhibited more elaborate explanations than individuals in their posttest. Arenz (1991) points out that although cooperative CAI is preferable to individual CAI, cooperative the former should not be considered as an absolute teaching method but rather as a supplement to regular instruction. Furthermore, Ahmad finds that individual and cooperative CAI must follow basic rules of environmental comfort such as accessibility of computers (Ahmad, Corbett, Rogers, & Sussez, 1985; Kubota, 1991):

> Microcomputers should not be put in a room where there is a lot of noise. Students should also have easy access to the machines if they are to be used outside normal hours. In other words, machines should be as straightforward and accommodating as possible. Access to tapes, discs and printed material must be easy and reliable. Lack of organization in this aspect can be annoying if students arrive to find the equipment already occupied or the disk unavailable (Ahmad, et al., 1985, p. 46).

Size of Computer Groups

Regarding the size of computer groups, interaction has been found to be comparable in dyads and triads (Guntermann & Tovar, 1987; Krendl & Lieberman, 1988). Other studies indicate that two was the optimal number (Fisher, 1984; Krendl & Lieberman, 1988): triads seem to compete among themselves, and when working in groups of four, students have more difficulty seeing the screen. Other studies indicate that in some cases, a whole class can work on a computer-based activity by enlarging its display on a larger screen with a projecting device (Ariew, 1990; Vockell, 1989): "In these cases, the computer can be viewed as a visiting teacher, who presents problems or ideas for the teacher and students to attack together" (Vockell, 1989, p. 31).

Overall studies investigating CAI and cooperative learning agree that cooperative computer-based activities can:

(1) pave the way to greater achievement;
(2) develop a better attitude and enhance the students' intrinsic motivation toward their studies, and
(3) make students actively involved in FL communication.

Post-achievement tests do not always show statistically significant differences in favor of cooperative computer-based activities. However, when tested in situations where higher cognitive reasoning skills are required (problem-solving situations, for instance) results show that cooperative computer-based activities provide good training (Guerrero, Eisler, Budin, & Lane, 1990; Johnson, et al., 1985).

Redefinition of CALL according to Recent Pedagogical Trends

What happens to foreign language interaction in the computer age? Since cooperative learning can enhance motivation and learning outcomes in foreign languages, and since computer-based group activities have shown to be beneficial in other academic areas, a combination of both cooperative learning and CALL has been recently used in foreign language-curricula:

> Mirroring recent changes in foreign language pedagogy, CALL is being redefined along the same lines. Rather than to propose more software that is repetitive and rigidly structured, new CALL software design entails giving students more flexibility, to seek and find information, to control not only the way they learn, but also to some extent the material they learn.
>
> For example, some new CALL materials are structured in such a way that students are given free reign to explore (or "browse") databases. Students are learning by discovery and by doing. This kind of freedom learning is very much reminiscent of some of the more

> modern classroom techniques such as group learning or student-centered learning (Ariew, 1991, p. 39).

CALL has been considered an asset for taking risks in language production and comprehension (Ariew & Frommer, 1987b). For instance, some word-processing programs provide ready access to language data such as a dictionary, a grammar, and the like. Students can explore language conventions by themselves without having to face red-ink corrections. When dealing with courseware, scores given at the end of activities or games are not definite and can be changed by doing the activity again. It allows L2 learners to accept errors or even failure as a normal learning process, not as a definite and threatening record (Morgan, 1986; Raschio, 1986), principles that fit the most recent constructionist views of education.

Such interactive flexibility with the computer allows the development of self-confidence which, according to Clément (1987) is a precondition for motivation. Likewise, CALL can encourage more risk-taking shown to positively predict language production (Ely, 1986a; Ely, 1986b). Keeping these assets in mind, a CALL environment seems to have the potential to provide the non-threatening context necessary for what Clément (1987) and Ely (1986a, 1986b) consider an essential element for the development of communicative competence.

Importance of Negotiations during Computer-Based Activities

Johnson (1991) mentions that beneficial interaction is not an automatic result of computer use but rather depends on the way language instructors "structure classroom interaction involving computer use" (Johnson, 1991, p. 62). For example, Piper (1986) conducted a study to investigate the type of target language that students in English as a Foreign Language (EFL) courses were generating at the computer among each other. The results indicated that in spite of the organization of the class in smaller groups, the type of software selected by instructors has a substantial impact on verbal interactions within groups. Piper observed that the selection of software can in fact counteract the benefits of a cooperative approach if it just involves structure manipulation.

Similarly, Mydlarski (1987) investigated the quality of language used among first-year French students who worked cooperatively with drill and game software. The results indicated that the more structured the courseware, the more frustrated students could become, and the more students tended to switch back to their native language. Drill courseware did not appear to encourage students to communicate naturally (conversations during such courseware are focused on spelling, grammar, vocabulary, or rule clarification). However, students relied more on group knowledge instead of using the on-line help option built into the courseware, which indicates that they naturally use cooperation.

Young (1988) conducted an experiment among intermediate-and advanced-level adult students of English to investigate differences in the type of discourse gener-

ated by a drill-and-practice program versus open-ended simulation software. Results indicated major differences between students working on negotiable outcome or non-negotiable outcome programs: "It was found that the negotiable-outcome programs generated discourse which was closer to ordinary conversation and entailed the kind of interaction which, a number of researchers have claimed, is beneficial to language acquisition" (Young, 1988, p. 65). Young's conclusions clearly show that it is not enough to gather students around a computer. Cooperative CALL depends on whether courseware allows students to negotiate the outcome of computer-based activities. Likewise, Degroff (1989) who investigated the role of computers in whole language classrooms, came to the conclusion that "for good teachers, it is the teacher's belief about curriculum and instruction rather than the technology that will determine the role of the computer in the classroom" (Degroff, 1989, p. 1).

Other studies have been conducted on the value of videodisk programs. Crotty (1984) compared two instructional settings: one class was traditionally conducted while instruction in the other class was assisted by a videodisk program identical in content with the material introduced in the traditional setting. Recall protocols indicated that the interactive videodisk group obtained superior results and showed a more positive attitude toward the material being introduced during instruction.

Likewise, so-called linear, segmented and Interactive Videodisk (IVD) programs (Verano, 1987) have been shown to produce different results. In Verano's study, students exposed to a linear program were passively viewing the material without interruption. Segmented programs were designed to present portions of the video with inserted comprehension check activities. The interactive program was similar to the segmented one except for the fact that students could replay the video, had access to a vocabulary list, and had extensive feedback (reinforcement of explanations in case of erroneous answers). Interactive videodisks were found to generate higher achievement scores than linear and segmented format.

Researchers in more recent studies (Chang, 1989; Chang & Smith, 1991) have investigated the effect of cooperative learning with an interactive videodisk program among beginning Spanish students who were either studying alone or in pairs at Interactive Videodisk workstations. The results of the study indicated that there was no statistically significant difference in learning between individual and cooperative learning at IVD workstations. Questions related to the video storyline, i.e. explicit and logical sequential items were answered equally well by individual and cooperative learners. However, answers to implicit items were more elaborate among cooperative learners, indicating that students working in groups had a better retrospective understanding of the story, due to a higher amount of discussion that took place within groups.

New Socio-Dynamism in and beyond Foreign Language Classrooms

Johnson (1991) recently stresses the fact that computers have become new participants in L2 instruction. As a result of computer use in the classrooms, socio-

dynamic factors (e.g. working together, peer teaching, friendship, positive atmosphere) affecting language acquisition have taken new shapes. Contrary to what was earlier feared, i.e. that computers were going to isolate students and negatively affect the efforts made in communicative language teaching, Johnson (1991) indicates that the CALL environment is, in many instances, promoting more authentic interactions which, as such, are more motivating than traditional instructional tasks:

> Rather than isolating and promoting asocial behavior, as many have feared, there is a growing body of evidence that computer use can promote new ways of working together, productive peer teaching, as well as high-quality social and academic task-based interaction, and that these kinds of interaction are related to higher levels of interest, motivation, and achievement. (Johnson, 1991, p. 65)

In an international experiment, Klobusicky-Mailander (1991) combined communication via computer network with team work . A simulation of the European Summit Conference held in Strasbourg was organized for universities across Europe and the USA. Although students of business, economics, politics, and English participated in the project were intimidated at first by the size of the project, they quickly became very involved and very motivated. From Karlsruhe where unified Germany's new role within the EEC was documented, the German team reacted to this large-scale simulation as follows:

> At the end of the simulation we discussed the results in our group and ascertained that we had gained enormously from the project both language-wise and as far as contents go. Our interest in current political and economic themes grew, whereby it was the game character of the simulation that increased our involvement in the individual topics. (Klobusicky-Maïländer, 1991, p. 50)

Similarly, Smith (1990) reports that the use of local area network and computer conferencing in foreign language instruction has proven to be beneficial for encouraging writing via modem among peers:

> Computer users improved significantly in their ability to read and express oral and written ideas. [...] interactive tools, whether word processing packages, computer conferences, or a combination of both, enable students to break their isolation, share ideas, and benefit from increased contact with teachers and peers. [...] Students who use computers write more and revise more, hence more extensive creative, and accurate production is possible. (Smith, 1990, pp. 80-81)

Other studies investigating synchronous computer networks used in L1 and L2 writing classes (Bump, 1990; DiMatteo, 1990; Kelm, 1992) indicate that CALL promotes substantially more written discussions and more involvement on the part of students. Observations of computer-assisted classes indicate that students are motivated by the fact that computer-based discussions appear to induce less language anxiety than writing formal papers, and thus students express their emotions and personal opinions more willingly. Generally, it has been noticed that with FL interactive activities within electronic networks, students become more involved and more motivated than their counterparts in more conventional settings (Bump, 1990; DiMatteo, 1990; Kelm, 1992; Klobusicky-Maïländer, 1991; Smith, 1990; Smith & Hoffman Maginis, 1990).

Immersion through Live-Action Simulations

Studies pertaining to the drawbacks in communicative language teaching have raised the issue of sociolinguistic authenticity within the limits of a classroom environment, which leads to the following question: How can authentic sociolinguistic situations be introduced into the instructional environment? The advent of interactive videodisks has raised a great deal of hope. Research in this area aims at bringing the real world into the classroom, and at providing the students with the sociolinguistic situations necessary for language use (Bush, Slaton, Verano, & Slayden, 1991). Today's interactive videodisks are designed to give students a sense of immersion through the use of videos and the reliance on a natural language processing system. Such systems exploit the strategies that take place in natural face-to-face conversations. Exchanges taking place with the computer become engaging while simultaneously enhancing believability. The operation of a video segment by the computer allows students to see, hear, and read language in context, a concept called "live action simulation:"

> This variety of language sources has the advantage of recreating the multidimensional nature of language as it exists within the visual, social, and cultural world, allowing for immersion in the truest sense of the word. Furthermore, the very nature of interactive video technology provides students with a degree of involvement heretofore unparalleled, allowing them to manipulate the video at will, and even to interact directly with characters in a story.
>
> [Today's IAV designs] exploit the immersion advantage to the fullest by "parachuting" the students into the middle of Paris or Bogotá where they are involved in a fictional story, with characters and a plot, and are also able to "wander" around. As a result, the learners find themselves in a completely authentic cultural and linguistic environment and come face to face with a very complex reality. Although this degree of authenticity and cultural density would present a clear problem with linear video, it is particularly well-

> suited to interactive video, since the computer can provide learners with all the tools they need to decipher and understand the whole cultural context in addition to the language. (Murray, Morgenstern, & Furstenberg, 1989, pp. 107-108)

These most recent trends in computer-assisted language learning indicate a bright outlook for future classroom interactions. If classrooms were equipped with only a small number of "desk-built-in platforms," it would not be hard to imagine the integration of "live-action-simulation group activities" in addition to text-based group activities that prevail today in FL instruction. Clearly, the presence of computers in the classroom will bring a much needed authentic environment that will open classrooms to the outside world. To date, however, most computer-based activities are conducted in language laboratories. Unless special arrangements are made to conduct a lesson in a computer laboratory, which is often done only occasionally due to the restricted availability of computer laboratories, computer-based group activities are rarely possible. Computers should be more accessible and not confined to a "special" place any longer. From the status of "special" tools, computers should be viewed as common and necessary tools to be integrated in foreign language classrooms. This article suggests that the computer laboratory should be expanded to foreign language classrooms to allow the integration of a much needed sociolinguistic authenticity to go beyond the restricting walls of the instructional environment. It is important to note that studies criticizing group work (Aston, 1986; Ellis, 1992a; Ellis, 1992b; Lightbown, 1992) fail to disconfirm the advantages of communicative language teaching, which is still believed to enhance confidence and motivation among L2 speakers. What these studies underline is the necessity to refine the practices of communicative activities, so as to enhance the quality of input. New techniques are thus needed to perfect communicative language teaching. In summary, computer-assisted language instruction seems to provide a promising answer.

Conclusion

Research reported in this article has shown that with the advent of computer-based communicative activities, the concepts of community, and sociolinguistic boundaries have taken other shapes for FL learners. Although a computer does not replace human interactions, nevertheless it fosters them. The following summarizes how CALL expands the sociolinguistic boundaries of the instructional environment and, as such, how it offsets the drawbacks of communicative language teaching.

Reported studies on FL interactions via computers show that students tend to feel less inhibited to share personal opinions and emotions while interacting with other members of a community, whether this community is linked to a local area network or to an international electronic network. Traditionally, classrooms have lent themselves to public talk, which does not always fit all personalities. Even if interventions are made less public with communicative activities, conventional classrooms

are still perceived as places where the instructor essentially monitors and structures interactions, circulating among groups to evaluate the students' participation. Although electronic interactions still generate public participation's, interestingly students do not feel directly watched or evaluated while using the language, which may explain why they are more willing to communicate via computers. Although not yet empirically researched, it seems that exchanging ideas via and around computers lends itself less to the evaluation of language forms than to the appreciation of messages and opinions. With this in mind, the advent of computers as communicative tools helps to:

(1) accommodate individual personality differences,
(2) reestablish meaningful, therefore authentic contacts, among students as well as between students and instructors,
(3) reach native speakers from all around the world via electronic networks.

Thus, the concept of authenticity in FL instruction needs to be redefined. Research reported in this article suggests that authenticity is not only a matter of natural language input provided by native speakers, but also a matter of natural speech and spontaneous exchanges among the members of a community, whether these exchanges take place between two non-native speakers or between a native speaker and a non-native speaker. Why should exchanges be labeled sociolinguistically authentic only when an element or a person of the target-language culture is present? Should this imply that the sociolinguistic reality of students is not valid? Would anybody disagree with the fact that inter-language exchanges are, in nature, as authentic sociolinguistic interactions as conversations taking place with a native speaker? Claiming that they are not implies that we keep looking at form instead of focusing on meaning. Clearly, computer-based activities seem to have been the impetus that helped reestablish the validity of meaningful communication among interlanguage speakers. Indeed, mention was made earlier in this article (Johnson, 1991) that open-ended computer-based activities foster natural communication on the side of students who, by their commitment, may very well involve instructors into their enthusiasm. What may have been partially perceived as lack of sociolinguistic authenticity in more conventional courses (Aston, 1986; Ellis, 1992; Lightbown, 1992) was probably due to the fact that the instructional environment still too often fails to foster meaningful exchanges. Even though done in an interlanguage, meaningful exchanges are still the foundation of an authentic discourse as well as the first sociolinguistic steps which help build enough confidence for exploring and sharing opinions on both local and international networks. The advent of computer-based conversations does not only encourage meaningful interlanguage exchanges on local area networks, it also creates opportunities to invite native speakers from international electronic networks to participate in classroom debates. Contrary to what was feared, i.e. that communicative activities can lead to a large-scale fossilization of erroneous language forms due to a mutual reinforcement of shared

interlanguage structures, today's CALL environment is promoting a quick access to high quality input by interacting electronically with native speakers. Clearly, computer-based interactions have expanded the sociolinguistic boundaries of the classroom to the outside world by integrating a large number of native speakers into the students' virtual connections.

Live-action simulations on interactive videodisks and CD-ROMs have also provided FL students with an interactive flexibility which has fostered the development of confidence. Indeed, research reported in this article has attested to the fact that CALL can encourage more risk-taking shown to predict positively language production and development. Even more attractive, live-action simulations have opened windows to the outside world. Today's interactive live action simulations give students a sense of immersion by allowing them to explore the entire planet. With the development of such simulations, computer-based language activities have become adventures: Students can choose to visit Paris, or—experience the wonders of Rio de Janeiro, or even explore the exoticism of Hong Kong. Within each language area, it is a matter of selecting the communicative software that best fits the adventurous appeal of the day. Just as we pull a book off the shelves to "travel" across time and space, we now select the software which takes us to the wonders not only of South America but also to the mysteries of Africa, to the beauties of Europe, and to the charm of Asia. Each of these computer trips takes students to culture-specific situations ranging from buying a plane ticket or ordering a drink to more challenging situations. Such FL simulations have been described as extremely stimulating and motivating to students: They "stimulate additional senses and involve the learner as closely as possible in experiencing the target language and culture without actually being present in a target language environment" (Schulz, 1991, p. 179). If students want to explore the computer land alone, they can be left alone; if they need assistance or if they have an extrovert personality, live-action-simulation-group-activities are wonderful ways to explore the target language culture with a group of friends. In either case, students are exposed both to high quality language input as well as stimulating cultural experiences.

Live-action-simulation group activities have an advantage over individual exploration. They foster interaction among students who end up negotiating to decide as a group what direction should be taken during their virtual visit. Thus, live-action-simulation group activities generate more language production than live-action simulations conducted individually. No medium can replace interaction among human beings. Live-action simulations are meant to stimulate students to explore the data of knowledge and to store it through situational associations. Such associations have been shown to increase memory (Asher, 1977) and to provide students with better procedural and declarative knowledge to use the language.

The case has been made here that CALL provides a rich source of interactive opportunities that expand the sociolinguistic boundaries of the communicative instructional environment beyond the limits of the classroom. As such, CALL not only promotes communication, it also brings the outside world to FL students who can

thus proceed with virtual explorations while being exposed to high quality language input.

References

Abraham, R. G., & Liou, H. C. (1991). Interaction generated by three computer programs: Analysis of functions of spoken language. In P. Dunkel (Ed.), *Computer-assisted language learning and testing: Research issues and practice.* New York: Newbury House.

Ahmad, K., Corbett, G., Rogers, M., & Sussez, R. (1985). *Computers, language learning and language teaching.* New York: Cambridge University Press.

Allwright, R. (1984). The importance of interaction in classroom language learning. *Applied Linguistics, 5,* 156-171.

Arenz, B. W. (1991). Mathematics problem solving within a cooperative learning computer structure. [EDRS: ED 334 971]

Ariew, R. (1990). CALL today: Implications for multisectioned language programs. In S. Magnan (Ed.), *Issues in language program direction.* Boston, MA: Heinle and Heinle.

Ariew, R. (1991). Effective strategies for implementing language training technologies. *Applied Language Learning, 2*(2), 31-44.

Ariew, R. & Frommer, J. G. (1987). Interaction in the computer age. In W. M. Rivers (Ed.), *Interactive language teaching.* New York: Cambridge University Press.

Asher, J. (1977). *Learning another language through actions: The complete teacher's guidebook.* Los Gatos, CA: Sky Oaks Productions.

Aston, G. (1986). Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics, 7*(2), 128-143.

Ballman, T. L. (1988). Is group work better than individual work for learning Spanish? The findings of one study. *Hispania, 71*(1), 180-185.

Breen, M. (1985). The social context for language learning: A neglected situation? *Studies in Second Language Acquisition, 7*(2), 135-158.

Brown, D. H. (1987). *Principles of language learning and teaching.* Englewood Cliffs, NJ: Prentice-Hall.

Brown, J. D. (1991). *Understanding research in second language learning.* New York: Cambridge University Press.

Bump, J. (1990). Radical changes in class discussions using networked computers. *Computers and the Humanities, 24,* 49-65.

Bush, M., Slaton, A., Verano, M., & Slayden, M. E. (Eds.). (1991). *Interactive videodisk: The "why" and the "how".* Provo, UT: Brigham Young University.

Bygate, M. (1988). Units of oral expression and language learning in small group interaction. *Applied Linguistics, 9*(1), 59-82.

Carrier, C. A. & Sales, G. C. (1987). Pair versus individual work on the acquisition of concepts in a computer-based instructional lesson. *Journal of Computer-Based Instruction, 14*(1), 11-16.

Chang, K.-Y. R. (1989). *The interaction of cooperative learning and computer-mediated interactive videodisk instruction in beginning Spanish.* Unpublished Dissertation, Purdue University.

Chang, K.-Y. R., & Smith, W. F. (1991). Cooperative learning and CALL/IVD in beginning Spanish: An experiment. *The Modern Language Journal, 75*(2), 205-211.

Chapelle, C., & Jamieson, J. (1986). Computer-assisted language learning as a predictor of success in acquiring English as a second language. *Tesol Quarterly, 20*(1), 27-46.

Clément, R. (1987). Second language proficiency and acculturation: An investigation of the effects of language status and individual characteristics. *Journal of Language and Social Psychology, 5,* 271-290.

Crotty, J. (1984). *Instruction via an intelligent videodisk system versus classroom instruction for beginning college French students: A comparative experiment.* Unpublished doctoral. dissertation, University of Kansas.

Davies, N. F. & Omberg, M. (1987). Pear group teaching and the composition class. *System, 15*(3), 313-323.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Degroff, L. (1989). Computers in the whole language classroom. *Florida Instructional Computing Conference.* Orlando, FL. [EDRS: ED 318 452]

DiMatteo, A. (1990). Under erasure: A theory of interactive writing in real time. *Computers and Composition, 7,* 71-89.

Di Pietro, R. J. (1987). *Strategic interaction: Learning languages through scenarios.* New York: Cambridge University Press.

Doughty, C., & Pica, T. (1986). "Information-gap" tasks: Do they facilitate second language acquisition? *TESOL Quarterly, 20*(2), 305-325.

Ellis, R. (1992a). The classroom context: An acquisition-rich or an acquisition-poor environment? In C. Kramsch, & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study.* Lexington, MA: DC Heath.

Ellis, R. (1992b). Learning to communicate in the classroom. *Studies in Second Language Acquisition, 14,* 1-23.

Ely, C. M. (1986a). Language learning motivation: A descriptive and causal analysis. *The Modern Language Journal, 70,* 28-35.

Ely, C. (1986b). An analysis of discomfort, risk taking, sociability, and motivation in the L2 classroom. *Language Learning, 36,* 1-25.

Faltis, C. (1990). Freirian and Vygotskian perspectives. *Foreign Language Annals, 23,* 117-126.

Feldmann, S. C., Fish, M. C., Friend, R., & Bastone, L. (1991). An observational study of social behavior in microcomputer classrooms. *Computers in Human Behavior, 7,* 45-55.

Fisher, G. (1984). The social effects of computers in education. *Electronic Learning,* 26-

28.

Freire, P. (1970). *Pedagogy of the oppressed.* New York: Continuum.

Gale, L. (1989). Macario, Montevidisco, and interactive dígame: Developing interactive video for language instruction. In F. Smith (Ed.), *Modern technology in foreign language education: Applications and projects.* Lincolnwood, IL: National Textbook.

Goodman, N. (1984). *Of mind and Other Matters.* Cambridge, MA: Harvard University Press.

Green, J. (1993). Student attitudes toward communicative and non-communicative activities: Do enjoyment and effectiveness go together? *The Modern Language Journal, 77*(1), 1-10.

Guerrero, F., Eisler, J., Budin, H., & Lane, P. (1990). Collaborative learning and computers: What some new York City teachers are doing. [EDRS: ED 320 542]

Guntermann, E., & Tovar, M. (1987). Collaborative problem-solving with LOGO: Effects of group size and group composition. *Journal of Educational Computing Research, 3*(3), 313-334.

Hartl, C. (1984). *Cooperation: Student pairs as a catalyst for learning Spanish.* Ph.D. Dissertation, Purdue University.

Higgs, T., & Clifford, R. (1982). The push toward communication. In T. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher. ACTFL Foreign Language Education Series.* Lincolnwood, IL: National Textbook.

Johnson, D. M. (1991). Second language and content learning with computers: Research in the role of social factors. In P. Dunkel (Ed.), *Computer-Assisted language learning and testing: Research issues and practice.* New York, NY: Newbury House.

Johnson, D. W., & Johnson, R. T. (1985). Cooperative learning: One key to computer-assisted learning. *The Computing Teacher, 13*(2), 11-15.

Johnson, D. W. (1986). Computer-assisted cooperative learning. *Educational Technology, 26*(1), 12-18.

Johnson, R. T., Johnson, D. W., & Stanne, M. B. (1985). Effects of cooperative, competitive, and individualistic goal structures on computer-assisted instruction. *Journal of Educational Psychology, 77*(6), 668-677.

Kagan, S. (1987). *Cooperative learning resources for teachers.* Riverside, CA: University of California.

Kelm, O. R. (1992). The use of synchronous computer networks in second language instruction: A primary report. *Foreign Language Annals, 25*(5), 441-454.

Klobusicky-Maïländer, E. (1991). Iconomics 101: A telecommunication simulation on Europe 1992. *Simulation/Games for Learning, 21*(1), 43-53.

Kramsch, C. J. (1985). Classroom interaction and discourse options. *Studies in Second Language Acquisition, 7*(2), 169-183.

Kramsch, C. J. (1987). Interactive discourse in small and large groups. In W. Rivers (Ed.), *Interactive language teaching.* New York: Cambridge University Press.

Krashen, S. D. (1985). *The input hypothesis.* London: Longman.

Krashen, S. D., & Terrell, T. D. (1983). *The Natural Approach: Language acquisition in the classroom.* Oxford: Pergamon.

Krendl, K. A., & Lieberman, D. A. (1988). Computers and learning: A review of recent research. *Journal of Educational Computing Research, 4*(1), 367-389.

Kubota, K. (1991). Applying a collaborative learning model to a course development project. *Annual Convention of the Association for Educational Communications and Technology.* Orlando, FL. [EDRS: ED 331 490]

Lightbown, P. M. (1992). Getting quality input in the second/foreign language classroom. In C. Kramsch, & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study.* Lexington, MA: DC Heath.

Littlewood, W. T. (1974). Communicative competence and grammatical accuracy in foreign language learning. *Educational Review, 27*(1), 34-44.

Long, M. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition.* Rowley, MA: Newbury House.

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207-228.

Mckenna, E. (1987). Preparing foreign language students to enter discourse communities in the U.S. *English for Specific Purposes, 6*(3), 187-202.

McLaughlin, B. (1987). *Theories of second language learning.* London: Edward Arnold.

Meravech, Z. R., Silber, O., & Fine, D. (1991). Learning with computers in small groups: Cognitive and affective outcomes. *Journal of Educational Computing Research, 7*(2), 233-243.

Meravech, Z. R., Stern, D., & Levita, I. (1987). To cooperate or not to cooperate in CAI: That is the question. *Journal of Educational Research, 80*(3), 164-167.

Meunier-Cinko, L. (1992). Interactive French language curricula of the future: A study of computer and video potential. *French Review, 66*(1), 147-153.

Morgan, L. Z. (1986). Computer-assisted instruction: How, when and why. *Foreign Language Annals, 19*(6), 515-518.

Murray, J. H., Morgenstern, D., & Furstenberg, G. (1989). The Athena language learning project: Design and issues for the next generation of computer-based language learning tool. In F. Smith (Ed.), *Modern technology in foreign language education: Applications and projects.* Lincolnwood, IL: National Textbook.

Mydlarski, D. (1987). Cooperative computer-assisted language learning: Is it living up to its promise? *Journal of Educational Techniques and Technologies, 20*(1), 26-29.

Omaggio-Hadley, A. (1993). *Teaching language in context (2nd ed.).* Boston, MA: Heinle and Heinle.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures.*

New York: Viking.

Pica, T. (1987). Second language acquisition, social interaction, and the classroom. *Applied Linguistics, 8*(1), 3-21.

Pica, T. & Doughty, C. (1985a). Input and interaction in the communicative language classroom: A comparison of teacher-fronted and group activities. In S. M. Gass, & C. Madden (Eds.), *Input in second language acquisition.* Rowley, MA: Newbury House.

Pica T. (1985b). The role of group work in classroom second language acquisition. *Studies in Second Language Acquisition, 7*(2), 233-248.

Pica, T., Holliday, L. & Lewis, N. (1991). Language learning through interaction. What role does gender play? *Studies in Second Language Acquisition, 13*, 343-376.

Piper, A. (1986). Conversation and the computer: A study of the conversational spin-off generated among learners of English as a foreign language working in groups. *Systems, 14*, 187-198.

Raschio, R. A. (1986). Communicative uses of the computer: Ideas and directions. *Foreign Language Annals, 19*(6), 507-514.

Reiser, R. (1987). Instructional technology: A history. In R. M. Gagne (Ed.), *Instructional technology: Foundations.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Rivers, W. (1983). *Communicating naturally in a second language. Theory and practice in language teaching.* New York, NY: Cambridge University Press.

Rivers, W. (Ed.). (1987). *Interactive language teaching.* New York: Cambridge University Press.

Saville-Troike, M. (1988). Private speech: Evidence for second language learning strategies during the 'silent' period. *Child Language, 15*, 567-590.

Scarcella, R. C., & Oxford, R. L. (1992). *The tapestry of language learning. The individual in the communicative classroom.* Boston: Heinle & Heinle.

Schachter, J. (1986). Three approaches to the study of input. *Language Learning, 36*, 211-226.

Schulz, R. (1991). Bridging the gap between teaching and learning. In S. S. Magnan (Ed.), *Challenges in the 1990s for college foreign language programs.* Boston, MA: Heinle & Heinle.

Seliger, H. W. (1983). Learner interaction in the classroom and its effects on language acquisition. In H. W. Seliger, & M. H. Long (Eds.), *Classroom-oriented research in second language acquisition.* Rowley, MA: Newburry House.

Sinclair, M., & Coulthard, M. (1975). *Towards an analysis of discourse.* Oxford: Oxford University Press.

Smith, K. (1990). Collaborative and interactive writing for increasing communication skills. *Hispania, 73*,77-86.

Smith, K., & Hoffman Maginis, B. (1990). VAX notes. *Computers and the Humanities, 24*, 146-153.

Smith, K. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-82.

Terrell, T. (1977). A natural approach to second language acquisition and learning. *The Modern Language Journal, 61*(7), 325-337.

Verano, M. (1987). *Achievement and retention in Spanish presented via videodisk in linear, segmented, and interactive modes*. Unpublished doctoral. dissertation, University of Texas.

Vockell, E. L. (1989). Group size for computerized instruction. *Educational Technology, 29*(9), 30-32.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Webb, N. M. (1984). Microcomputer learning in small groups: Cognitive requirements and group processes. *Journal of Educational Psychology, 76*(6), 1076-1088.

Webb, N. M. (1989) Peer interaction and learning in small groups. *International Research, 13*(1), 21-40.

Webb, N. M., & Kenderski, C. M. (1984). Student interaction and learning in small-group and whole-class settings. In P. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes*. Orlando, FL: Academic Press.

Webb, N. M., Ender, P., & Lewis, S. (1986). Problem-solving strategies and group processes in small groups learning computer programming. *American Educational Research Journal, 23*, 243-261.

Winn, W. D. (1989). Toward a rationale and theoretical basis for educational technology. *Educational Technology, Research and Development, 37*(1), 35-46.

Wong-Fillmore, L. W. (1992). Learning a language from learners. In C. Kramsch, & McConnel-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study*. Lexington, MA: DC Heath.

Young, R. (1988). Computer-assisted language learning conversation: Negotiating an outcome. *CALICO Journal, 5*(3), 65-83.

Author

LYDIE E. MEUNIER, Assistant Professor, Department of Languages, The University of Tulsa, 600 South College Avenue, Tulsa, OK, 74104. Specialization: second language acquisition and teaching.

Applied Language Learning
1994, Vol. 5, No. 2, pp. 57-93

FL/L2 Reading, Grammatical Information, and the Input-to-Intake Phenomenon

David J. Shook
Georgia Institute of Technology

Research has shown that second language L2 learners have a limited capacity for processing the input to which they are exposed. The present study focused on the input-to-intake phenomenon: Can foreign language FL/L2 learner-readers process grammatical information presented via written input as intake? Reading passages were designed which varied the amount of attention drawn to the Spanish present perfect and the relative pronouns que/quien(es). *Multiple tasks were designed to assess the processing of input into intake by first- and second-year students of Spanish. Results indicated that only when FL/L2 learner-readers have their attention explicitly drawn to the grammatical items will they process them as intake; in addition, they process the input for more-meaningful structures before processing for less-meaningful ones. First- and second-year subjects exhibit about the same amount of gain in grammatical information; however, their ability to recognize or to produce the grammatical items might depend on their language experience, the meaningfulness of the items involved, and the passage-dependency of the task.*

Introduction

Whereas language acquisition traditionally has been described in linguistic terms, focusing on the inherent properties of languages themselves, interest has grown recently in examining language acquisition within the frameworks of information processing and general learning as seen in terms of cognitive psychology (O'Malley & Chamot, 1990). Researchers in cognitive psychology view human beings as limited-capacity processors who cannot process all of the information they are exposed to at one time due to task demands and the individual's experience with such information, among other factors (McLaughlin, Rossman, & McLeod, 1983). That foreign- and second-language (FL/L2) learners only take in some and not all of the language to which they are exposed also can be attributed to this limited processing capacity. In addition to the FL/L2 task demands (e.g. searching the input for vocabulary items on a listening

© 1994, David J. Shook

comprehension test) and experience (e.g. the similarity of new verb forms in the input to those already learned, for example); factors such as content, grammatical complexity, and pragmatic information will determine what part of the language is processed. FL/L2 researchers such as VanPatten (1985, 1989) have advanced these ideas to examine what part of the language to which a learner is exposed, the *input*, actually is processed by the learner in some way, the *intake*. Intake, in this paper, is defined in terms of what Slobin (1979) and Peters (1985) refer to as language *extracted/segmented* by the learner, or that part of the language input that Tomlin and Villa (1994) would label as *detected* by the learner; questions of awareness, consciousness, and control aside, intake is that part of the language input to which the learner is exposed that he/she selects for further processing. In sum, when one studies those factors contributing to whether or not certain aspects of language are attended to by the language learner, one has begun to examine the *input-to-intake phenomenon*.

VanPatten states that, from an input-processing point of view, there are three processes involved in FL/L2 development. During Process I, certain structures in the input are singled out to become the learner's intake (*input-to-intake*). In Process II, those structures made available during Process I are further singled out so that they then become part of the learner's repertoire of acquired structures (*intake-to-acquisition*). Finally, during Process III, the language learner selects some message to communicate from the available FL/L2 structures (*acquisition-to-use*). Thus, a possible scenario regarding an FL/L2 learner is the following (from VanPatten and Cadierno, 1993):

I II III

INPUT —> INTAKE —> developing system —> output

In the above diagram, as one moves from left to right, the words are printed smaller and smaller. Thus, this graphically represents the limited capacity of language learners to process language: Due to internal and external processing restraints, only part of the input becomes intake, only part of the intake becomes acquired, and only part of the acquired intake can be used by the learner.

Therefore all stages of the language-learning process, input-to-intake, intake-to-acquisition, and acquisition-to-use, are interdependent: by definition, if part of the input does not become intake, it can not become acquired by the learner, and certainly that part of the language cannot be used by that learner. Therefore, in order for researchers in the field of FL/L2 development to describe accurately not only what a learner has developed in terms of language ability but also how he/she developed it, research needs to examine each stage of the process from an input-processing perspective. Consequently, the logical first step is to investigate VanPatten's Process I: What input becomes intake and how does this occur?

Previous Research

Three areas of study have insights for this line of research: studies of syn-

thetic-grammar processing, FL/L2 oral-input processing, and FL/L2 written-input processing. The first important studies on specific linguistic features in the input-to-intake phenomenon were undertaken in the domain of cognitive psychology. A multitude of studies have investigated learning in general and language acquisition in particular, by exposing subjects to synthetic (artificial) grammars that would eliminate natural-language effects due to specific languages or language families. Although debate exists as to the nature of explicit/unconscious learning versus implicit/conscious learning of synthetic grammars, and how the assessment tasks utilized affect of such learning (see, for examples, the excellent reviews in Segar [1994] and Shanks & St. John [1994]), some synthetic grammar studies do provide evidence regarding the extraction and the segmentation abilities of L2 learners. Reber (1976), Dulany, Carlson, and Dewey (1984), Reber, Regan, and Allen (1985), as well as Morgan, Meier, and Newport (1987) have examined the effects of the input from synthetic grammars on learners' intake.

Reber (1976) and Morgan et al. (1987) suggest that learners develop knowledge of grammatical rules, specifically those regarding constituent structure, through exposure to artificial grammars under various processing conditions (implicitly and explicitly forming rules; presence and lack of clues to phrasal constituency). While Dulany et al.(1984) state that their attempt to replicate Reber's study produced contrary results regarding the implicit versus explicit conditions necessary for learning, Reber et al.'s (1985) reassessment of Dulany et al.'s research design and data suggests that Dulany et al.'s study does provide evidence that learners have the ability to extract and segment grammatical rules from the input. In sum, there is evidence that grammatical information in the input found in synthetic grammars is available to be processed as intake. However, although all of these studies utilized similar post-exposure tasks to determine intake, the results offer conflicting evidence regarding whether implicit processing or explicit processing leads to better results on those types of tasks.

The findings of Zobl (1985) and VanPatten (1990) suggest that, as was the case with synthetic grammars, learners are able to process some grammatical information found in the FL/L2 oral input as intake. In the case of Zobl's study, the evidence suggests that learners were able to process grammatical information pertaining to the markedness of the possessive determiners *his/her* in English. VanPatten's study gives evidence that learners can process different types of grammatical information (lexical items, articles, and verb morphemes), but that attention to such information impacts negatively on the processing of content.

Berman (1984), Lee (1987), and Shook (1989) all examined the comprehension and processing of grammatical structures presented via written input. These studies suggest that information such as English syntax (Berman), the Spanish subjunctive (Lee), and the Spanish present perfect (Shook) can be processed by FL/L2 learners through written input, and not necessarily interfere with the successful comprehension of the input. However, none of these studies offers any evidence for the input-to-intake phenomenon. While Berman and Lee only attempted to demonstrate what text information was or was not comprehended, and Shook did attempt to examine whether or not input was processed as intake, methodological problems with the post-exposure

assessment tasks in Shook's study limit the conclusions that can be drawn from two of the three assessment tasks (see Shook [1989,1990] for more details).

Leow (1993) did investigate the input-to-intake phenomenon in the FL/L2 reading process. He specifically examined the effects of simplification, type of linguistic item, and language experience on learners' intake of linguistic items found in written input. First- and second-year learners of Spanish at the university level were exposed to one of four conditions: a simplified or unsimplified reading passage involving the present perfect verb tense, or a simplified or unsimplified passage involving the present subjunctive. Results were based on the analyses of subjects' performance on a pre- and post-exposure multiple-choice recognition task. Leow found significant main effects for passage and language experience, and a significant interaction between passage and task: however, the subjects who read the simplified passages significantly increased their recognition scores from pre- to post-test in comparison to their counterparts who read the unsimplified passages; in addition, second-year subjects significantly attended more to the linguistic items in the input than did the first-year subjects. Leow concludes that simplification of the FL/L2 written input did not make a facilitating effect on learners' intake of linguistic items from the input; although it may affect comprehension, simplification of the FL/L2 input may not facilitate "those processes that aid intake" (1993, p. 343). However, as FL/L2 linguistic development increases, it does seem that the ability to take in more linguistic information also increases.

Summary and Limitations of the Previous Research

A comparison of the studies previously cited is found in Table 1. Although there exists research into the processing of written input, the majority of the studies have not examined the input-to-intake phenomenon. In general, it has been shown that in the studies that investigated input processing by means of exposure to synthetic grammars (Reber, 1976; Dulany et al., 1984; Morgan et al., 1987), and FL/L2 oral input (Zobl, 1985; VanPatten, 1990), the learners were able to process some grammatical information from the input. However, Leow (1993) appears to be the only published researcher that addressed successfully the issue of intake from FL/L2 written input.

Of the studies reviewed above, only VanPatten (1990) and Leow (1993) investigated the processing of more than one grammatical feature of the input. They are the only researchers cited who can make any statements regarding the processing of different grammatical features under the same experimental conditions.

Although similar in their use of varying exposure conditions, in the previously-cited studies the types of exposure conditions utilized were not uniform. To add consistency to the findings on input processing, researchers should attempt to employ multiple-exposure conditions to the input. One of the first ways to do this would be to focus on varying the amount of attention drawn to grammatical items present in the input.

TABLE 1
Previous Studies of Processing Grammatical Information

	Input to Intake?	Number of Grammatical Structures	Number and Type of Exposure Conditions	Number and Type of Post-Exposure Assessment Tasks	Tasks Dependent on Input Content?	Multiple Levels of Language Learners?
Reber (1976)	yes	1	2 instructions to memorize input; instructions to figure out rules	1 grammaticality judgment	dependent only	no
Dulany et al. (1984)	yes	1	2 instructions to memorize input; instructions to figure out rules	1 grammaticality judgment	dependent only	no
Morgan et al. (1987)	yes	1	3 no clues for phrasal groupings; incorrect clues; correct clues	2 recognition; grammaticality judgment	dependent only	no

Zobl (1985)	yes	1	2 marked vs. unmarked input	1 written answers to oral questions	dependent only	no
VanPatten (1990)	yes	2	4 attention to passage content; attention to content + verb morpheme; attention to content + article; attention to content + lexical item	1 free written recall	dependent only	3 levels: 1st semester; 4th semester; 3rd year
Berman (1984)	no	1	2 simplified vs. unsimplified input	2 multiple-choice questions; open-ended questions	dependent only	no

SSL

730

Lee (1987)	no	1	1	5 cloze passage; probe questions (English); probe questions (Spanish) free written recall (English); free written recall (Spanish)	dependent only	2 levels 1st semester 2nd semester
Shook (1989)	no	1	3 varied amounts of attention called to grammatical items	2 free written recall; multiple-choice recognition	dependent and independent	no
Leow (1993)	yes	2	2 simplified vs. unsimplified input	1 multiple-choice recognition	dependent only	2 levels 1st semester 4th semester

A majority of the studies cited employed only one type of post-exposure task to assess the information about grammatical rules that was taken in: only the FL/L2 written-input studies were consistent in the use of multiple post-exposure assessment tasks. The need for multiple post-exposure assessment tasks should be obvious; the more tasks employed, the more confidence there is in the results achieved. However, another aspect of the use of multiple post-exposure assessment tasks needs to be addressed. Of all the tasks employed in the research cited, the majority of the studies used "recognition" tasks, that is, tasks that only required subjects to recognize or identify the grammatical structures to which they were exposed, such as grammaticality judgment and multiple-choice questions. The other studies all have used tasks that could be classified as "production" tasks, requiring that the subjects produce (usually by writing) some or all of the grammatical information to which they were exposed. Thus, it is apparent that no input study has yet employed both recognition and production post-exposure assessment tasks within the same research design. Finally, after exposure to grammatical input, are subjects limited to using that input (in recognition tasks, in production tasks, etc.) in situations that are dependent on the topic content of the input, or can subjects use the input in situations that are independent of the topic content of the input? Whether FL/L2 learners are limited in their ability to use the knowledge gained in the input-to-intake phenomenon, and whether or not this knowledge is constrained to the topic content of the original input, needs to be examined.

Of the studies related to input/intake, only VanPatten (1990), Lee (1987), and Leow (1993) included more than one level of language learner in their studies. The inclusion of multiple levels of subjects would broaden insights into the input-to-intake phenomenon. Studies that treat multiple levels of subjects would probe the accessibility of certain grammatical features for different levels of language learners and would allow a more accurate picture of the developmental characteristics of the input-to-intake phenomenon, that is, how more experience with the language might affect the ability to process the input.

This review of the previous research has highlighted certain needs in this area of research. They call for more FL/L2 written input studies that (1) go beyond information processing to examine the input-to-intake phenomenon; (2) study the processing of more than one grammatical feature; (3) vary the attention conditions under which subjects are exposed to the input; (4) employ multiple post-exposure assessment tasks, including both recognition and production tasks; and (5) examine multiple levels of language learners. The remainder of this article reports on a study that addressed the above-mentioned needs within a single-research design.

Research Questions and Variables

The following research questions guided the present study:

(1) What effect does explicitly drawing attention to grammatical items in the input have on FL/L2 learner-readers' intake?

(2) What effect does the type of grammatical item in the input have

on FL/L2 learner-readers' intake?
(3) What effect does target-language experience have on FL/L2 learner-readers' intake?

To bridge some of the gaps mentioned in the previous research and examine whether or not FL/L2 learner-readers can process grammatical information, presented via written input, as intake, the following variables were examined:

Attention Conditions

Under what conditions are grammatical items in the input more or less accessible to FL/L2 learners? In the present study, "attention" was defined as "focusing on, noticing, having one's attention drawn to, being conscious of" the grammatical items in the input (cf. Schmidt, 1990; VanPatten, 1990). One independent variable in the present study was the manner in which the input was enhanced, in other words, how the conditions under which the subjects focus on the grammatical items were varied (cf. Sharwood-Smith, 1988; 1991).This variable consisted of three levels: no attention drawn to the items, attention drawn to the items, and attention drawn to the items plus focus on forming grammatical rules.

Grammatical Features

Are all grammatical features of the input equally accessible to FL/L2 learners as intake? Two different grammatical items in Spanish were chosen for the present study: The present perfect verb tense and the relative pronouns *que/quien(es)*. (Examples of the grammatical structures utilized in this study and the contexts in which they ocurred are found in Appendices A and B.)

One aspect that English-speaking learners of Spanish have to learn is how to express a perfective action, i.e., one that has taken place immediately prior to the point of orientation.The use of the present perfect in Spanish is due to a choice on what *meaning* needs to be expressed: an action that is seen as completed (Ramsey & Spaulding, 1956); one that expresses "anteriority to the recalled axis of orientation" (Bull, 1965, p. 154). (While the present perfect has other uses in Spanish, and while other verb tenses in Spanish can suggest a perfective action, these aspects are not addressed in the present study.) The present perfect in Spanish is formed by a conjugated form of the verb *haber* plus the appropriate past participle. In the present study, only the forms of the present perfect for the third person, singular and plural, were examined.

Another aspect that English-speaking learners of Spanish have to learn is how to express "who/whom" in relative clauses: with *que* or *quien(es)*? The choice is a syntactic one. The Spanish relative pronoun *que* is used for [+/- human] antecedent noun phrases; in restrictive relative clauses, when the antecedent is animate, *que* cannot follow a preposition, but in nonrestrictive relative clauses, *que* can alternate with *quien(es)*, even after prepositions. On the other hand, *quien(es)* must be used with a [+human] antecendent after a preposition in restrictive relative clauses (Liceras, 1986).

For example,

(1) El hombre *que* acaba de llegar organizó el viaje.
(The man *that/who* just arrived organized the trip.)
(2) Ese hombre, *que/quien* lleva traje, es guapo.
(That man, *who* wears a suit, is handsome.)
(3) Son las chicas *sobre quienes* se escribió el artículo.
(They are the girls *about whom* the article was written.)

The two structures differ primarily in that the use of the present perfect is determined by a meaningful, aspectual decision on the part of the speaker/writer, while the use of the relative pronouns *que/quien(es)* in restrictive clauses is determined by a syntactic decision; due to this distinction, for the present study, the present perfect is deemed to carry more meaning than the relative pronoun. Is the more-meaningful present perfect easier to process (e.g. take in from the input, extract and/or segment) than the less-meaningful relative pronoun?

Language Experience

Do learners with different amounts of the FL/L2 exposure and experience process the input differently? Another independent variable of this study was "language experience," consisting of two levels of learners at the university level: first year and second year.

Task Type

What post-exposure tasks best assess the processing of input into intake? If the FL/L2 learner has attended to the input, what has the learner developed through the input-to-intake process? Two possibilities involve being able to identify proper contexts of usage of a grammatical feature (a "recognition" ability) and being able to produce the grammatical features in the correct context (a "production" ability). The present study employed multiple post-exposure assessment tasks: recognition and production tasks for the present perfect and the relative pronouns.

Task Content

Another issue that is raised relates to the situations in which the learner-readers are able to recognize and/or produce the information gained. Are FL/L2 learner-readers limited in recognizing/producing the present perfect and the relative pronouns to tasks that reflect the topic and linguistic content of the written input (passage-dependent content), or are they also able to recognize/produce those grammatical structures in tasks that do not reflect the topic and linguistic content of the written input (passage-independent content)? In addition to the four types of tasks designed, each type of task

included two types of topic content: (1) content dependent on the passage input and (2) content independent of the passage input.

Methodology

Subjects

The subjects were 125 undergraduate students in their first or second year of Spanish at the university level; 60 enrolled in the first-year course and 65 in the second-year course. These groups were represented by subjects from Spanish 122 (first semester) and Spanish 124 (fourth semester) from the University of Illinois. There were 78 male and 47 female students with a mean age of 20.19 years. The subjects represented the four first years of college study. All but six subjects reported some study of Spanish in high school: first-year subjects reported a mean of 1.9 years and second-year subjects a mean of 2.13 years of high school study of Spanish. In addition, 21 of the subjects reported that they had studied other languages for periods ranging from one semester to more than four years, with French being the other language most often studied. Both subjects who indicated on a post-test background questionnaire that they remembered studying the grammatical items previous to the present experiment and those who did not indicate any previous study of the items were included in the subject pool. This was done to investigate whether subjects gained more information after reading the input passages in relation to what information they had beforehand.

Reading Passages

Two different reading passages were chosen to provide the written input which presented the grammatical features to the subjects, one passage for each grammatical item. The passage chosen for the present perfect segment of this study was adapted from the article "Fuera de órbita" from *Semana* (10 January 1989), a Colombian magazine (see Appendix A). The passage discusses the discovery of chaotic planetary orbits in our solar system. The article was chosen because of the natural occurrence of the Spanish present perfect in a number of sentences. The article was rewritten to include 3 more tokens of the verb tense in question; this conversion required changing tense of verbs from the preterit to the present perfect. Afterwards, there was a total of 6 instances of present perfect verbs: 3 each for third person singular and third person plural. The total length of the article was shortened by eliminating some secondary information. In its final form, the passage contained 185 words. The passage was considered typical of the type of reading passages utilized in the basic language program in Spanish at the university.

The passage chosen for the relative pronoun segment of this study was adapted from the article "Jim Thorpe, el sioux que asombró al mundo" from *El País Semanal*, the Sunday supplement to *El País*, a Madrid newspaper (see Appendix A). The passage was one in a series of articles that appeared in the magazine in anticipation of the 1988 Olympic Games in Seoul. The "Thorpe" passage contained natural occurrences of the

relative pronoun *que*. The original article was rewritten to convert tokens of *que* with inanimate antecedents to instances with [+human] antecedents and to others with [+human] antecedents followed by a preposition and *quien(es)*; in total, there were three tokens of the relative pronoun *que* and three of *quien(es)* with [+human] antecedents in the rewritten passage. The passage was also shortened by eliminating secondary information; in the end, there were 217 words in the passage. This article also was deemed typical of the reading passages utilized in the basic language program at the university.

In conjunction with the passages, a free written recall of each passage was used to focus the subjects on reading for comprehension. The recall was written in English, following Lee (1986a & 1986b), who suggests that recalls done in the native language of the learner allow subjects to demonstrate to a better degree their comprehension of an FL/L2 text. The use of the recall for this study was to give a sense of realism and purpose to the reading task: the subjects were instructed to read to comprehend the passage information (see Appendix A). Although the results of the recall task were not analyzed for this study, the data they provide will be included in a future report.

Assessment Tasks

A series of assessment tasks was designed to elicit subjects' recognition and production abilities of both grammatical structures after exposure to the input passages (for examples of the tasks, see Appendix B). There were four types of tasks:

(1) Production of the Present Perfect

This section consisted of 12 cloze sentences to test whether or not the subjects could produce the present perfect in situations that called for its use. The subjects were required to write the correct present perfect form of a verb, given in its infinitive, to complete the form/function requirements of a Spanish sentence. The 6 items that were dependent on the content in the passage required that the subjects conjugate the verb exactly as it appeared in the passage. Six additional items also tested the production of the present perfect, but were not dependent on the passage input.

(2) Recognition of the Forms and Uses of the Present Perfect

This section consisted of 12 multiple-choice sentence completions in which subjects had to choose the correct missing verbal elements. Six of the questions related information from the passage, with the present perfect verb forms listed as a possible answer among four choices, and 6 additional items were not dependent on the information from the passage.

(3) Production of the Relative Pronouns *que/quien(es)*

For the 12 items in this section, subjects were asked to supply the appropriate word to

correctly complete the sense of the sentence. Six items required subjects to complete sentences which repeated material from the passage with the relative pronoun input; another 6 items were not dependent on the passage information, but still required the use of relative pronouns.

(4) Recognition of the Relative Pronouns *que/quien(es)*

Again, there were 12 items in this multiple-choice, sentence-completion task. Subjects had to choose the correct missing phrase out of four distractors. Six of the phrases reflected information from the passage, with *que* and *quien(es)* listed as possible answers among four choices, and an additional six items were passage independent.

The "information gained" about the grammatical input provided by the two reading passages was operationalized as the subjects' scores on post-exposure assessment tasks in comparison to pre-exposure assessment tasks. The pre-test included the same items that were utilized in the post-exposure tasks. The first four sections of the pre-test alternated between six relative-pronoun and six present-perfect production tasks. The passage-dependent and passage-independent items were randomly distributed within each section. Distractors were included in each section so that the subjects could not guess the underlying purpose of the experiment. The fifth section of the pretest focused on the recognition of both the relative pronouns and the present perfect within the same section. This section consisted of 30 items: the twelve post-exposure tasks for each grammatical structure (presented in random order) plus six distractors.

Testing Procedure

All testing was carried out in the subjects' regular classroom during their regularly scheduled class meetings, and was administered by the present experimenter in 10 of the 18 sections. Three class periods were necessary for the testing procedure: The second day of testing occurred two days after the first, and the third day occurred four days after the second. On each day, subjects were allowed 40 minutes to complete the required tasks. (Only during the first test period did the subjects utilize most of the allotted time; in the two remaining test periods, the subjects required between 20 and 35 minutes to complete the input passage and the post-exposure test.) On the first day of testing, each subject received a testing booklet which presented the pre-exposure assessment tasks to be performed. The five pre-exposure test sections were presented in the order mentioned above, production tasks preceding recognition tasks, to avoid any additional attention to the grammatical input that might have occurred if the recognition tasks had preceded the production tasks. On the second and third days of testing, each subject received a testing booklet which presented the tasks to be performed. The order of tasks was as follows:

(1) Read warm-up passage for the day (counterbalanced in presentation).
(2) Recall warm-up passage.
(3) Read input passage for the day (Counterbalanced in presentation: half of the subjects received the present perfect passage on the second day; the other half received the relative pronoun passage on the second day).
(4) Recall testing passage.
(5) Post-exposure assessment tasks for the day (production tasks first, recognition tasks second).
(6) Debriefing questionnaire on the grammatical item (subjects identify whether or not the item had been studied before).

For each group of subjects, the only difference in the materials was found in the written input to which they were exposed. For the control group (condition 1), there was no change made to the passage and no special instructions were given. For the second group (condition 2) the six tokens of the grammatical item were printed in bold, uppercase letters, but there were no special instructions nor was subjects' attention explicitly drawn to those items. Finally, the third group (condition 3), also received the reading passage with the grammatical items in bold and uppercase letters (see Appendix A). However, in addition to the instructions to read the passage and to understand the information, the following instructions were included for condition 3:

> At the same time, notice the words that appear in BOLD, UPPERCASE LETTERS. Try to come up with a rule for the use of those particular words.

To control for group and order effects, the testing materials were randomly distributed to the test subjects in each class section so that equal numbers of the subjects received each of the three exposure conditions.

Scoring Procedure

Due to factors such as subject absences and non-compliance with instructions, not all of the students who participated in the testing procedure formed part of the final subject pool. The following subject groupings resulted for analysis: 35 subjects received the "condition 1" passage (18 first-year and 17 second-year subjects), 52 received the "condition 2" passage (23 first-year and 29 second-year subjects), and 38 subjects received the "condition 3" passage (19 first-year and 19 second-year subjects). The first step in the analysis of the results was the scoring of the pre- and post-test items, followed by the computation of the gain scores. Gain scores are an effective and accepted means of testing for the effect of some experimental treatment when the same pre-and post-testing tasks are utilized (Kern, 1989; Scott, 1989). Gain scores are

achieved by subtracting the pre-test scores from the post-test scores. A positive gain score would indicate a positive effect from exposure to the input; a zero gain score would indicate no effect (either positive or negative); finally, a negative gain score would indicate a negative effect.

The following criteria were utilized in order to compute the gain scores for each subject. On the present perfect production tasks, both the correct form of *haber* plus the correct past participle were necessary for complete credit; anything else was deemed incorrect and no credit was given. On the relative pronoun production tasks, the correct *que* or *quien* was necessary for complete credit; no credit was given for anything else. On the recognition tasks, again either complete or no credit was given: for both grammatical items, the correct choice out of the four possibilities received full credit—everything else received no credit. The subjects' mean gain scores were used as the values for the dependent variables in tests of significance.

Results

In this section, overall scores on the experimental tasks are provided first to arrive at a general understanding regarding the subjects' performance. Next, there follow the results of the statistical procedures utilized to explore the effects of the various variables examined. Finally, this section provides the distinct cell comparisons needed to describe the findings revealed through the statistical procedures.

Overall Scores

Both subjects who indicated on the post-test questionnaire that they had or had not studied the grammatical items previously comprised the subject sample of this study. Twenty-five (41.6%) of the 60 first-year subjects indicated that they had not studied the present-perfect previously; 16 (24.6%) of the 65 second-year subjects indicated the same. For the relative pronoun, 29 (48.3%) and 21 (47.7%) of the first- and second-year subjects, respectively, indicated no previous study of that grammatical item. The present perfect was clearly more familiar to both the first- and second-year subjects, while almost half of both groups of subjects were not familiar with the relative pronouns. More second-year subjects were familiar with both items than were first-year subjects.

Overall mean gain scores by the 125 subjects for each task type are presented in *Table 2*. (All gain scores have a possible minimum value of -6 and a possible maximum value of 6, and the following values are computed from the scores in Table 2.) Both the highest and the lowest mean gain scores came on the content-dependent recognition tasks: 1.800 for the present perfect tasks and 0.208 for the relative pronoun tasks. Taking all of the tasks into account, the overall mean gain score was 1.073, with a range from -5 to 6. For all of the present perfect items alone, the range was -5 to 6 with a mean gain of 1.318; for the relative pronoun items, the range was -2 to 6 with a mean gain of 0.828. On the production items alone, the gain scores ranged from -2 to 6 with a mean of 1.278, while the recognition scores ranged from -5 to 6 with a mean

of 0.866. Finally, the content-dependent items alone exhibit a range of -5 to 6 with a mean of 1.264, while the content-independent items also ranged from -5 to 6, but exhibited a mean gain of 0.882. In sum, it is evident that the scores on the various tasks do not vary to a large degree. While the mean gain scores per condition are not very large, there is still evidence for a positive effect for exposure to the input passages.

TABLE 2
Mean Gain Scores by Item, Task, and Content

Item	Task	Content	Score	SD
Present Perfect	Production	Dependent	1.400	2.178
Present Perfect	Production	Independent	0.976	1.657
Present Perfect	Recognition	Dependent	1.800	2.163
Present Perfect	Recognition	Independent	1.096	2.030
Relative Pronoun	Production	Dependent	1.656	1.805
Relative Pronoun	Production	Independent	1.088	1.581
Relative Pronoun	Recognition	Dependent	0.208	1.727
Relative Pronoun	Recognition	Independent	0.368	1.609

Inferential Statistical Analysis

The particular procedure used in this study was a repeated-measures MANOVA design. Since all the subjects were exposed to both of the grammatical items examined and all subjects took all of the same pre- and post-exposure tests for each item, the repeated measure was the variable *grammatical item*. The search for variance in the MANOVA procedure differentiates those factors which separated the subjects into separate groups ("between-subjects effects") from those factors which involved all subjects ("within-subject effects"). In the present case the between-subjects effects were the variables *attention condition* and *language experience*, while the within-subjects effect was the variable *grammatical item*.

Three different MANOVA procedures were conducted in connection with the experimental data, and Tables 3-5 present these data. The first was a 3 X 2 X 2 X 4 MANOVA involving the three levels of exposure condition, the two levels of language experience, the two grammatical items, and the gain scores on four tasks: production with content-dependent items, production with content-independent items, recognition with content-dependent items, and recognition with content-independent items (Table 3). Over all the test items, significant main effects were found for exposure condition ($p \leq 0.001$) and for grammatical item ($p \leq 0.01$), as well as a significant interaction between grammatical item and language experience ($p \leq 0.01$).

TABLE 3
MANOVA Summary for Overall Gain Scores

Between-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	1692.47	238	7.11			
CO	258.09	2	129.05	18.15	0.000	***
EX	1.38	1	1.38	0.19	0.660	n.s.
CO by EX	11.03	2	5.52	0.78	0.462	n.s.

Within-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	1459.19	714	2.04			
IT	50.83	1	50.83	7.15	0.008	**
CO by IT	11.36	2	5.68	0.80	0.451	n.s.
EX by IT	49.77	1	49.77	7.00	0.009	**
CO by EX by IT	4.52	2	2.26	0.32	0.728	n.s.

Notes:
CO = Attention Condition
EX = Language Experience
IT = Grammatical Item
** = $p \leq 0.01$, *** = $p \leq 0.001$

TABLE 4
MANOVA Summary for "Type of Task" Gain Scores

Between-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	3648.94	244	14.95			
CO	516.19	2	258.09	17.26	0.000	***
EX	2.76	1	2.76	0.18	0.688	n.s.
CO by EX	22.07	2	11.03	0.74	0.479	n.s.

Within-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	1727.32	244	7.08			
IT	79.43	1	79.43	0.23	0.001	***
CO by IT	3.21	2	1.61	0.23	0.797	n.s.
EX by IT	27.94	1	27.91	3.95	0.048	*
CO by EX by IT	13.98	2	6.99	0.99	0.374	n.s.

Notes:

CO = Attention Condition
EX = Language Experience
IT = Grammatical Item
* = $p \leq 0.1$, *** = $p \leq 0.001$

TABLE 5
MANOVA summary for "Type of Content" Gain Scores

Between-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	3648.94	244	14.95			
CO	516.19	2	258.09	17.26	0.000	***
EX	2.76	1	2.76	0.18	0.668	n.s.
CO by EX	22.07	2	11.03	0.74	0.479	n.s.

Within-Subjects Effects

Source of Variation	Sum of Squares	df	Mean Square	F	p	Level of Sig.
Within Cells	915.78	244	3.75			
IT	72.28	1	72.28	19.26	0.000	***
CO by IT	0.01	2	0.01	0.00	0.998	n.s.
EX by IT	0.99	1	0.99	0.26	0.608	n.s.
CO by EX by IT	1.33	2	0.66	0.18	0.838	n.s.

Notes:

CO = Attention Condition
EX = Language Experience
IT = Grammatical Item
*** = $p \leq 0.001$

The second procedure was a 3 X 2 X 2 X 2 MANOVA, involving the same independent variables as the previous MANOVA, but utilizing only the dependent variable of *type of task gain score* (see Table 4). Upon examination of only the data from type of task, that is, production vs. recognition tasks, the same results were revealed: significant main effects for exposure condition ($p \leq 0.001$) and for grammatical items ($p \leq 0.001$) and a significant interaction between language experience and grammatical item ($p \leq 0.05$).

The final procedure was another 3 X 2 X 2 X 2 MANOVA, again including the same independent variables as before, but this time utilizing the dependent variable of *type of content gain score* (see Table 5). This procedure only found significant main effects for exposure condition ($p \leq 0.001$) and grammatical item ($p \leq 0.001$).

Cell Comparisons

Scores for the passage-dependent items were higher for both production and recognition tasks than were scores for the passage-independent items (see Table 6). In addition, production scores were higher than recognition scores, regardless of content. Matched t-tests confirmed that production scores were significantly different from recognition scores (observed t = 1.822; for 249 df and a probablility level of 0.05, critical t = 1.658), and that content-dependent scores were significantly different from content-independent scores (observed t = 2.248; for 249 df and a probablility level of 0.05, critical t = 1.658).

TABLE 6
Mean Gain Scores: Task and Content

		Content	
		Dependent	Independent
Task	Overall	1.264	0.882
Production	1.279	1.528	1.032
Recognition	0.866	1.000	0.732

TABLE 7
Mean Gain Scores: Attention Condition

Task	Content	Condition		
		1	2	3
Production	Dependent	0.7143	1.6923	2.0526
Production	Independent	0.2143	1.1154	1.6711
Recognition	Dependent	0.3143	1.0192	1.6053
Recognition	Independent	0.0571	0.8365	1.2105

As seen in Table 7, gain scores for condition 3 (attention called to the items plus focus on forming rules) were consistently higher than the gain scores for condition 2 (attention called to the items), which were consistently higher than the gain scores for condition 1 (no attention called to the items), regardless of type of task or content. Although the MANOVA procedures revealed a significant main effect for exposure condition, they do not by themselves reveal the nature of that effect when there are more than two levels to a variable. Therefore, Tukey post-hoc tests were run for each of the experimental tasks under the exposure condition variable. A significant difference in gain scores only came about between conditions that differentiated between "no attention called to the grammatical items in the input" (condition 1) and those that called for "attention to the grammatical items" (conditions 2 and 3). A differentiation between amounts of attention (conditions 2 and 3) was not statistically significant.

Although the MANOVA procedures revealed a main effect for grammatical item, closer examination of the data reveal some interesting patterns to the data. Table 8 presents the mean gain scores achieved on the experimental tasks for the two grammatical items examined. Regardless of the task content (context-dependent or context-independent), the gain scores on the present perfect items are shown to be consistently superior to those for the relative pronoun items. However, between the item and task variables, subjects performed slightly better on the relative pronoun items on the production tasks; on the recognition tasks, subjects performed much better on the present perfect items. It must be remembered that any main effect must be interpreted in the light of any significant interactions. The MANOVA procedures revealed a significant interaction between grammatical item and language experience, but only for the production/recognition tasks. Therefore, it is necessary to examine the data concerning that interaction before any specific conclusions are reached (see Table 9).

TABLE 8
Mean Gain Scores: Grammatical Item

Item	Task Production	Recognition	Content Dependent	Independent
Present Perfect	1.1880	1.4930	1.6000	1.0360
Relative Pronoun	1.3720	0.2084	0.9280	0.7280

TABLE 9
Mean Gain Scores: Interaction of Grammatical Item and Language Experience on Task

Production Task Experience	Present Perfect	Relative Pronoun
First Year	0.7250	1.6417
Second Year	1.6154	1.1231
Recognition Task **Experience**	**Present Perfect**	**Relative Pronoun**
First Year	1.4500	0.5417
Second Year	1.4462	0.0462
First Year	1.4500	0.5417
Second Year	1.4462	0.0462

Table 9 shows the scores for grammatical item and language experience. Given that there is an interaction effect for these two variables, there is no one level of either that is consistently superior/inferior to any other level of the variables. However, three patterns do exist: (1) On the present perfect production tasks, the second-year subjects did better than the first-year subjects. (2) On the relative pronoun production tasks, first-year subjects did better than second-year subjects. (3) On the recognition tasks, both first-year and second-year subjects did better on the items treating the present perfect than they did on the relative pronoun items. The subjects' performance on the production tasks depended on the level of language experience: first-year did better producing the relative pronouns and second-year did better producing the present perfect. On the other hand, the subjects' performance on the recognition tasks depended on the grammatical item: first-and second-year both did better recognizing the present perfect than they did the relative pronouns.

Discussion

Overall, the subjects in the present study gained linguistic information about the grammatical items of the study after exposure to the input passages, although such gains were relatively small. In continuation, all of the variables manipulated in the study and their respective research questions are addressed.

> *What effect does explicit drawing of attention to grammatical items in the input have on FL/L2 learner-readers' intake?*

A significant main effect was found for the attention condition across all of the data, indicating that subjects whose attention was drawn to the grammatical items gained more linguistic information about the items than the subjects whose attention was not called to the items. Also, the type of attention called to the input was shown not to be a significant factor. Better performance on the tasks after reading the input passages indicates that the subjects were producing/recognizing more instances of the present perfect and the relative pronouns. Those subjects whose attention was drawn to the grammatical items in the input performed better. How might these results be interpreted? One interpretation is related to the issue of Operating Principles (OP's).

Both Slobin (1979) and Peters (1985) have proposed the use of OP's on the part of L1 learners. Such OP's allow the learners to store salient chunks of the input (extraction) and later analyze these chunks (segmentation) as they build up their knowledge of the grammatical rules of their L1. For example,

> EX:COMPARE. Determine whether a newly extracted chunk of speech seems to be the same as or different from anything you have already stored (Peters, 1985, p. 1033).
> SG:STRESS. Segment off a stressed syllable of an extracted unit and store it separately (Peters, 1985, p. 1038).

VanPatten (1989, 1990) has suggested that FL/L2 learners might take advantage of similar rules. The results of the present study regarding attention conditions to the input may provide further insights into the possibility of FL/L2 learners using OP's. These might be called "Operating Principles of Attention." Since there was a significant difference between attention condition 1 and attention conditions 2 and 3, but not between 2 and 3, the following might be an Attention OP for FL/L2 learners:

> AT:ATTEND. Pay attention to and notice anything that sticks out in any way in speech and writing.

This would help learners notice grammar and vocabulary which, in writing, is highlighted, glossed, appears in different colors, etc., and which, in speech, is emphasized, has a new or different sound, etc. Such attention might allow learner-readers to take in information about grammatical features found in FL/L2 reading when the readers' attention is focused on such features, or they "stick out" in some way, better than when no such attention is drawn to the items. Tomlin and Villa (1994) developed a model of the role of attention in second language acquisition. Their model incorporates three components: alertness, orientation, and detection. While they theorize that only detection is necessary for acquisition, alertness and orientation may enhance detection (p. 197). The results of the present study support such a model of attention in that the learner-readers detected information regarding the present perfect and the relative pronouns in Spanish found in written input, and such detection was enhanced by instructions to focus on the highlighted grammatical features in the input. With this in mind, a new version of the Attention OP proposed above may be formulated:

> AT:DETECT. Pay attention to and notice anything that sticks out in any way in speech and writing, then select it for further processing.

In sum, the explicit drawing of attention to grammatical items in the input seems to help FL/L2 learner-readers detect those items, allowing them to become intake.

> *What effect does the type of grammatical item in the input have on FL/L2 learner-readers' intake?*
> *What effect does target-language experience have on FL/L2 learner-readers' intake?*

A significant main effect was found for grammatical item across all conditions; in addition, a significant interaction was found between grammatical item and language experience, overall and for type of task. These results indicate that regarding task, the subjects did better on the present perfect tasks than they did on the relative pronoun tasks. However, while the ability to produce the grammatical items depended on the type of item and the level of experience of the subject, recognition knowledge

gained depended only on the type of grammatical item.

Positive mean gain scores overall for both grammatical items suggest that the subjects did gain information regarding the grammatical items after exposure to both input passages, and therefore it is possible that this new information became intake for the subjects. However, the significant main effect for grammatical item indicates that for the most part, more was taken in about the present perfect than about the relative pronoun. But on production tasks, relative pronoun gains were higher, and on recognition tasks, present perfect scores were higher. What might cause these differences? The answer might lie in the relative meaningfulness of the grammatical items.

The use of *que* or *quien* is a purely syntactic decision, governed by the type of relative clause (restrictive or non-restrictive) and the presence/absence of prepositions; the use of the present perfect, on the other hand, is an aspectual choice, determined by the meaning to be expressed. After processing the grammatical input found in the reading passages, the more meaningful the structure, the easier it was for the subjects to recognize when to use it in the appropriate situation. As VanPatten (1985) proposed, language learners might rely on two different but not mutually exclusive processing strategies. At first, all conscious attention is focused on processing *meaning*; since all conscious attention is occupied, no attention is available for the processing of *linguistic form*. However, once processing for meaning becomes automatic, then and only then does conscious attention become available for processing form. Along the same lines, then, if language learners are processing for meaning first and then processing for form, it would seem logical that more-meaningful linguistic structures would be processed before less-meaningful ones. Therefore, in the present study, the subjects took in more about the more-meaningful feature of the input (the present perfect) than they did for the less-meaningful feature (the relative pronoun); this pattern is supported by the learners' performance on the recognition tasks.

However, the results on the production tasks do not reflect the same patterns: gain scores were higher for the relative pronoun items. This phenomenon appears to be related to the language experience of the language learners.

No significant main effect for language experience was found in any of the MANOVA procedures. This means that neither first-year nor second-year subjects took in relatively more information about the grammatical features of the input than did the other group. However, a significant interaction between grammatical item and language experience was found overall and for the type of task variable. Second-year subjects performed significantly better on the present perfect production tasks than did first-year subjects; but first-year subjects performed significantly better on the relative pronoun production tasks than did second-year subjects. The fact that the second-year subjects performed better on the production tasks related to the present perfect should be expected: they had more experience and study of the FL/L2, and thus should do better than the first-year subjects. However, this did not prove true regarding the relative pronoun production tasks: first-year subjects performed better than did second-year subjects. What might explain these results?

Post-hoc analyses have eliminated two possible explanations. Did the second-

year subjects achieve higher pre-test scores of the relative pronoun production tasks, due to having more experience with the FL/L2, than did the first-year subjects, and thus achieve relatively lower gain scores after exposure to the input passages? The pre-test scores showed a mean score of 1.325 for the first-year subjects and 1.569 for the second-year subjects on the relative pronoun production tasks, which proved not to be significantly different enough to support this interpretation of the results. A second possible explanation might have been found in examination of what the subjects actually wrote during the production tasks. It might have been the case that the second-year subjects, with more language experience, had available a larger inventory of grammatical structures to call upon, and thus tried to use the relatives *el cual, la cual, los cuales, las cuales* and other pronouns instead of *que/quien*. Again, however, a post-hoc analysis of the relative pronoun production tasks did not support this conclusion. First-year subjects used *que/quien* about 70 percent of the time after reading the input passages, while second-year subjects used those relative pronouns about 66 percent of the time. No substantial use by either groups of subjects was found for the forms of *cual* or for any other pronouns. Therefore, further analysis of the subjects' performance on the relative pronoun production tasks is necessary before a definitive conclusion is reached.

Finally, overall, subjects performed significantly better on production tasks than they did on recognition tasks, and better on the tasks which were content-dependent than on those which were content-independent. While mean gain scores were positive for both recognition and production tasks, why were production scores higher than recognition scores? This would seem to be due to the very low gains by subjects in recognizing when to use the relative pronouns. As seen in Table 7 above, both first- and second-year subjects failed to achieve gains of much more than a mean of .05 items. Subjects tended to score high on the pre-test, then perform much worse on the post-exposure tasks. In addition, as discussed above, the results suggest that it was easier for the subjects to recognize when to use a more meaningful structure than a less meaningful one, and therefore recognition (good or bad) was a function of grammatical structure. However, production of the grammatical items depended on the language ability of the subjects. Both groups of subjects achieved mean gains of more than 1.5 for one particular grammatical item. While it seems logical for the second-year subjects to do better on the present perfect production tasks than the first-year subjects, first-year subjects' better performance on the relative pronoun production tasks requires further analysis for adequate explanation.

Although overall mean gain scores were positive for both content-dependent and content-independent items, the subjects found it easier to produce and recognize the present perfect and the relative pronouns when completing tasks with the themes of "chaotic planetary orbits" and "Jim Thorpe," but they found it more difficult to produce and use those same grammatical items in novel, appropriate situations. This should not be surprising. In order to recognize/produce the grammatical information processed in situations independent of the task content, the function of the intake would have to be acquired in some manner. This implies storage procedures not in evidence in Process I, but which are more appropriate to Processes II and III (see discussion below).

What grammatical information was taken in by the learner-readers? Due to the fact that all of the subjects were early-stage language learners (first- and second-year), who tend to use lower-level/bottom-up processing strategies (cf. the review of the research literature found in Bernhardt [1991]), it seems evident that the subjects in the present study processed the majority of the information in the reading passages at the level of the sentence, and not at the level of the paragraphs or the passage as a whole. Therefore, clues towards meaning and form were sought within each individual sentence. For example, regarding the present perfect, there were adverbials present (*últimamente, desde los primeros años de observaciones*, etc.) that provided clues to the perfect aspect of the verb. On the other hand, the relative pronoun *que* followed a [+human] antecedent; however, if a preposition followed the [+human] antecedent, *quien* was present. Such clues also were present sentences utilized in the pre- and post-exposure tasks. This arrangement indicates that the grammatical information detected by the learner-readers while reading included the fact that one particular structure (the present perfect) often appears when there are words giving a "perfective aspect"—however that aspect might be defined by early-stage learners. In the same manner, the grammatical information tapped might have included the fact that another structure (the relative pronoun que) is not used "when a preposition is present"—again, however this rule might be defined by early-stage learners.

It is noteworthy that VanPatten's model of input processing by FL/L2 learners highlights one of the assumptions that guided the present study: namely, that not all of the input to which a language learner is exposed is actually taken in by the learner. The three independent variables which were manipulated in the present study were selected in order to examine the role of input processing with regards to the grammatical features found in FL/L2 reading passages. As Figure 1 represents, these variables can be integrated into VanPatten's model to highlight their usefulness in understanding the processing of grammatical features of the FL/L2 input. Thus, the grammatical features comprised the input to which the FL/L2 learner-readers' attention was drawn to varying degrees. The attention conditions under which the learner-readers received the grammatical input can be seen as a factor which may affect Process I. In addition, the language experience of the learner-readers also can be seen as an influential factor of the input-to-intake phenomenon. The tasks which assess whether or not the learner-readers could recognize the grammatical features of the input would relate to the product of Process II, the internalized FL/L2 system of the learner-reader. The tasks which assess whether or not the learner-readers could produce the grammatical features of the input would relate to the product of Process III, that is, the FL/L2 learner-readers' output. The differentiation between content-dependent and content-independent tasks allowed for further insights into the breadth of the learner-readers' recognition and production abilities.

FIGURE 1 *Factors Affecting the Processing Grammatical Features of the FL/L2 Input.*

Note: Based on VanPatten, B. & T. Cadierno (1993a). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15(2), 225-244.*

760

760

The following question might arise from the preceding discussion: How might it be ascertained that the present study did tap the input-to-intake phenomenon (Process I), as was its intention, and not the input-to-intake-to-acquisition phenomenon (Process II), or the input-to-intake-to-acquisition-to-use phenomenon (Process III)? To reflect accurately either Process II or Process III, the study would have had to account for the storage of the input; that is, whatever language is detected in Process I needs to be stored in some way, most probably in long-term memory. In the present study, there was no real opportunity for long-term memory storage of language data, since the exposure to the written input and then the assessment tasks ocurred within a 40-minute time span. It is most improbable that the data collection procedures used could reflect anything except the immediacy of Process I, and thus this study does not reflect any acquisition of the grammatical input.

Therefore, the present study offers evidence about the nature of the input-to-intake phenomenon. The results suggest that information about grammatical features of the input can be taken in/processed/detected by early-stage FL/L2 learners when those grammatical items are presented via written input, that is, via reading passages. This ability to process the grammatical input increases when attention is drawn to the grammatical items in the input. In addition, it seems that the more meaningful the grammatical item, the more successful the learners are at being able to recognize when to use the item. However, the amount of experience with certain FL/L2 structures and/or with the FL/L2 overall may bring about differences in the ability of the subjects to produce the grammatical items. Finally, these recognition and production abilities can manifest themselves in novel situations, although it seems to be easier for the learners to recognize and produce the items with information related to the input passages.

Future Research

The present study should not be seen as a one-time project; it does not stand alone, but rather it builds upon previous research and, at the same time, suggests avenues for further research. One area of future research concerns the role of attention to the input in the input-to-intake phenomenon. The present study examined three different levels of noticing the input (cf. Schmidt, 1990), or of input enhancement (cf. Sharwood-Smith, 1991). Would different means of drawing learners' attention conditions (instructions without highlighted forms, different highlighting techniques, etc.) bring about differences in processing the grammatical input? This study provides evidence supporting Tomlin and Villa's (1994) model of attention in second language acquisition. However, further research in this vein is needed to substantiate this model for FL/L2 development.

Another important aspect of the present study is that its findings are limited to only two grammatical features of Spanish: the present perfect and the relative pronoun *que/quien(es)*. How is the processing of these grammatical features of Spanish by FL/L2 learner-readers different from or similar to the processing of other grammatical

features of Spanish? Are the abilities of early-stage FL/L2 learner-readers to process linguistic information about the present perfect and relative pronouns via written input generalizable to all early-stage learner-readers of other FL/L2s? If this is evidence of some universal input-processing strategy, how does it relate to L1 early-stage Operating Principles such as those suggested by Slobin (1979) and Peters (1985)?

Another limitation of the present study concerns the experimental subjects. First- and second-year learner-readers of Spanish at the university level comprised the subject pool for this experiment, and tended to process the grammatical information in similar ways. How might the input-to-intake phenomenon differ between other groups of subjects, for example, intermediate, advanced, and native-like learner-readers? In addition, both subjects who had studied the grammatical structures and those who had not participated in the study. Given that the present study has provided evidence that early-stage FL/L2 learners in general can get more information about grammatical structures from reading passages, it would be logical for future research into the input-to-intake phenomenon to differentiate between those subjects with and those without previous experience with the input.

An additional avenue for further research relates to the issue of post-exposure assessment tasks. The present study utilized production tasks (fill-in-the-blank sentences) and recognition tasks (multiple-choice cloze sentences); in addition, each type of task involved both passage-dependent and passage-independent items. What other tasks might provide more and/or different information about the process of taking in grammatical information via written input?

One final limitation of the present study regards the use of gain scores for comparing the effects of the various variables examined. The present study utilized a strict criterion of acceptability for determining the gain scores for each subject. What more might be revealed about the input-to-intake phenomenon if a loose criterion of acceptability or a point system designed to give partial credit for subjects' responses was adopted for the scoring procedure?

Pedagogical Implications

The role that the explicit teaching of grammar should or should not play in the FL/L2 classroom is a debate that has waged for many years within the profession. The results of the present study provide some food for thought for FL/L2 teachers and FL/L2 researchers who are interested in this issue.

One point to consider is that all grammatical items are not processed alike. The processing of certain items as intake may require explicit attention to be drawn to the grammatical items in the input, whereas the extraction/segmentation of other items may not. Therefore, while FL/L2 acquisition researchers must try to determine what structures are processed in similar or different ways and why this occurs, FL/L2 instructors must not assume that an effective means for providing learners with access to one grammatical item for processing will necessarily be effective for a different type of structure, for different levels of learners, or for different purposes.

FL/L2 acquisition researchers such as Ellis (1984), Lightbown and d'Anglejan

(1985), and Sharwood-Smith (1988, 1991) have discussed the research on the effects of formal instruction on FL/L2 acquisition, and all agree that formal instruction, while probably not influencing the natural route of FL/L2 acquisition to a great extent, might affect the *rate* of FL/L2 acquisition by providing shortcuts for learning or by providing access to comprehensible input that might not be accessed in other ways. Both Ellis (1984) and Sharwood-Smith (1988) agree that "formal instruction" should not be interpreted simply as "focus on form," but rather in terms of "elaboration" and "explicitness," as seen in Figure 2.

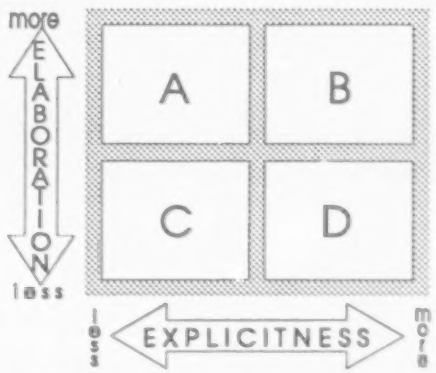


FIGURE 2. *Degrees of Formal Instruction.*

Note: From Sharwood-Smith, M. (1988). Consciousness raising and the second language learner. In W.. Rutherford & M. Sharwood-Smith (Eds). *Grammar and second language teaching* (pp. 51-60). New York: Newbury House.

Taking the present study as an example, if certain reading passages can provide learner-readers with the access to certain aspects and or types of grammatical input that they need, without explicit attention drawn to the grammatical input (for example, boxes A or C in Figure 2), perhaps this will allow instructors to save more explicit teaching (boxes B or D) for those grammatical items needing that explicit attention (see VanPatten & Cadierno [1993a, 1993b] for further discussion). In addition, reading passages selected to "instruct" on some grammatical point can be tailored to fit into the appropriate levels of explicitness and elaboration needed to allow learner-readers access to some grammatical input. Thus, reading passages may be useful to FL/L2 instructors not only as sources of grammatical input, but also as tools for tailoring formal instruction to the needs of the FL/L2 learners.

Finally, the subjects in the present study did not make large, substantial gains in information/abilities after their exposure to the grammatical features via written in-

put. Again, this should not be surprising: the FL/L2 acquisition process is naturally a building-up process, comprised of extracting meaningful chunks of language from the input and segmenting those chunks into more manageable pieces, in the course of which new information about the grammatical structure of the language is gained. In addition, more meaningful structures tend to be processed before less meaningful ones. Therefore, FL/L2 instructors should not be disenchanted by seemingly slow gains in their students' acquisition of linguistic/grammatical information. Rather, they should focus on the *direction* of the gain instead of on the *magnitude* of the gain. In addition, increased explicit attention to the grammatical features found in the input should not be seen as necessarily the most effective instructional method for FL/L2 acquisition to occur. Given the natural pace of linguistic information gain, such a methodology might prove frustrating to both instructor and learner. However, as Krashen (1982, and elsewhere) has advocated, a classroom environment that is rich in comprehensible input is a desireable factor for FL/L2 development in general. In this same vein, increased exposure to FL/L2 input as a whole through reading passages will provide the learners with more opportunities to process the input they need; a natural by-product of this increased exposure to FL/L2 input in general will be increased exposure to the grammatical input, which, depending on the attention drawn to the grammatical input, may be used as a short-cut for FL/L2 learning.

Conclusion

New information about the input-to-intake phenomenon has been revealed through the present study. While previous research has shown that language learners have the ability to access grammatical input by means of exposure to synthetic grammars and FL/L2 oral input, the present study offers evidence that language learners can process grammatical features of the input by means of exposure to written input. In addition, the present study suggests that factors such as attention to the input, level of the language learner, and meaningfulness of grammatical feature all may affect the ability of the learner to take in the grammatical information found in the written input.

References

Berman, R. A. (1984). Syntactic components of the foreign language reading process. In J. C. Alderson & A. H. Urquhart (Eds.). *Reading in a foreign language* (pp. 139-156). New York: Longman.

Bernhardt, E. B. (1991). *Reading development in a second language*. Norwood, N. J.: Ablex Publishing Corporation.

Bull, W. E. (1965). *Spanish for teachers*. New York: Ronald Press.

Dulany, D., Carlson, R. & Dewey, G. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 11*(3), 541-555.

Ellis, R. (1984). *Classroom second language development*. Oxford: Pergamon Press.

Kern, R. (1989). Second language reading strategy instruction: Its effects on compre-

he[illegible] and word inference ability. *Modern Language Journal, 73(2)*, 135-149.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. New York: Pergamon Press.

Lightbown, P. M. & d'Anglejan, A. (1985). Some input considerations for word order in French L1 and L2 acquisition. In S. M. Gass & C. G. Madden (Eds.). *Input in second language acquisition* (pp. 415-430). Rowley, MA: Newbury House.

Lee, J. F. (1986a). On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition, 8*(2), 201-212.

Lee, J. F. (1986b). The effects of three components of background knowledge on second language reading. *Modern Language Journal, 70*(4), 350-354.

Lee, J. F. (1987). Comprehending the Spanish subjunctive: An information processing perspective. *Modern Language Journal, 71*(1), 50-57.

Leow, R. P. (1993). To simplify or not to simplify: A look at intake. *Studies in Second Language Acquisition, 15*(3), 333-356.

Liceras, J. (1986). *Linguistic theory and second language acquisition: The Spanish nonnative grammar of English speakers*. Tübingen: Gunter Narr Verlag.

McLaughlin, B. (1987). *Theories of second-language learning*. London: Edward Arnold.

McLaughlin, B., Rossman, T. & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*(2), 135-158.

Morgan, J. L., Meier, R. P. & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrase to the acquisition of language. *Cognitive Psychology, 19*, 498-550.

O'Mally, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Peters, A. M. (1985). Language segmentation: Operating principles for the perception and analysis of language. In D. I. Slobin (Ed.). *The cross-linguistic study of language acquisition: Theoretical issues* (Vol. 2 , pp. 1029-1067). Hillsdale, NJ: Erlbaum.

Ramsey, M. M. & Spaulding, R. K. (1954). *A textbook of modern Spanish*. New York: Holt, Rinehart and Winston.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2*(1), 88-94.

Reber, A. S., Allen, R. & Regan, S. (1985). Syntactical learning and judgement, still unconscious and still abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General, 114*, 17-24.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 127-158.

Scott, V. M. (1989). An empirical study of explicit and implicit teaching strategies in French. *Modern Language Journal, 73*(1), 14-22.

Segar, C. A. (1994). Implicit learning. *Psychological Bulletin, 115*(2), 163-194.

Shanks, D. R., & St.John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367-447.

Sharwood-Smith, M. (1988). Consciousness raising and the second language learner. In W. Rutherford & M. Sharwood-Smith (Eds.). *Grammar and second language teaching* (pp. 51-60). New York: Newbury House.

Sharwood-Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*, 118-132.

Shook, D. J. (1989). *Can learners get grammatical intake from the input?* Paper presented at the 1989 Second Annual SIP SLATE Colloquium. Urbana, IL.

Shook, D. J. (1990). *The processing of grammatical structures via written input*. Unpublished doctoral dissertation, University of Illinois, Urbana/Champaign.

Slobin, D. I. (1979). *Psycholinguistics*. Glenview, IL: Scott & Foresman.

Tomlin, R. S. & Villa, V. (1996). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-203.

VanPatten, B. (1985). Communicative value and information processing in L2 acquisition. In E. Judd, P. Nelson, & D. Messerchmitt (Eds.). *On TESOL '84: A brave new world* (pp. 89-99). Washington, DC: TESOL.

VanPatten, B. (1989). *Towards a model of input processing*. Paper presented at the 1989 American Association of Teachers of Spanish and Portuguese, 71st Annual Meeting, San Antonio.

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*, 287-301.

VanPatten, B. & Cadierno, T.(1993a). Explicit instruction and input processing. *Studies in Second Language Acquisition,15*(2), 225-244.

VanPatten, B. & Cadierno, T.(1993b). Input processing and second language acquisition: A role for instruction. *Modern Language Journal, 77*(1), 45-57.

Zobl, H. (1985). Grammars in search of input and intake. In S. M. Gass & C. G. Madden (Eds.). *Input in second language acquisition* (pp. 329-349). Rowley, MA: Newbury House.

Appendix A

Sample Reading Passages
Passage for the present perfect: Condition 3
(emphasized forms and instructions)

Instructions: Read the following article through so that you understand the information presented. You will be asked to recall the information from the article, so focus on comprehending the passage. At the same time, notice the words that appear in **BOLD, UPPERCASE LETTERS**. Try to come up with a rule for the use of those particular words. When you have finished, turn to the next page.

Fuera de órbita

Durante años, desde el descubrimiento de Plutón hasta el presente, se **HA DISCUTIDO** acerca de la estabilidad del sistema solar. La interacción de los nueve planetas y el hecho de que éstos no se choquen, soprende hasta los científicos más expertos. El mismo Newton, después de proponer la ley de gravedad, declaró que una continua intervención divina **HA SIDO** indispensable para asegurar el órden del sistema.

Recientemente, en el laboratorio del Instituto Tecnológico de Massachusetts, los científicos Gerald J. Sussman y Jack Wisdom se **HAN ENTERADO** de que la órbita del planeta Plutón es caótica. Por medio de complicadísimos cálculos determinaron el movimiento de los cinco planetas durante 845 millones de años. Utilizando una computadora digital, los científicos ahora **HAN ENCONTRADO** que la órbita de Plutón es caótica en un período de 20 millones de años.

Desde los primeros años de observaciones, el científico Sussman **HA PENSADO** que todo el sistema solar es caótico y ahora tiene evidencia que sí es así. Debido a los recientes descubrimientos, los científicos **HAN TENIDO** que cambiar sus ideas originales en cuanto al universo. Puede ser que Newton después de todo tenía razón.

Note: The same exact passage was used for conditions 1 and 2, except that for both of these conditions there appeared no instructions to focus on the specific grammatical features, and for condition 1, the present perfect verbs were not highlighted in any way.

Passage for the Relative Pronouns: Condition 3
(emphasized forms and instructions)

Instructions: Read the following article through so that you understand the information presented. You will be asked to recall the information from the article, so focus on comprehending the passage. At the same time, notice the words that appear in **BOLD, UPPERCASE LETTERS**. Try to come up with a rule for the use of those particular words. When you have finished, turn to the next page.

Jim Thorpe, el sioux **QUE** asombró al mundo

El silencio y la humillación no derrumbaron la leyenda de Jim Thorpe, el sioux sobre **QUIEN** habló todo el mundo durante los Juegos Olímpicos de 1912. Su nombre siempre se asociará a su actuación en las pruebas de pentatlón y decatlón de los Juegos en Estocolmo. Thorpe consiguió las dos medallas de oro y estableció un récord mundial de decatlón; el récord perduró 17 años.

Sin embargo, dos meses después de ganar los dos títulos, un periódico publicó una fotografía de jugadores de béisbol con **QUIENES** apareció Thorpe, enfundado en

la camiseta de un equipo semiprofesional. Su participación en aquella Liga de verano vulneró los rígidos principios del amateurismo. Nada impidió el castigo del Comité Olímpico Internacional: suprimió el nombre de Thorpe de las listas de campeones olímpicos y le retiró las medallas.

Thorpe vivió sus últimos años abandonado y pobre, ignorado por los oficiales **QUE** borraron su nombre y marcas de las tablas de récordes. Murió Thorpe en la miseria en 1952. Tuvieron que pasar 32 años para que el COI devolviera las medallas de oro a su hija **QUE** las recibió en vísperas de los Juegos de 1984. Por fin, se restablecieron el nombre y las marcas en los libros registros olímpicos del hombre de **QUIEN** se comentó que era "el jugador teóricamente perfecto" y el mejor atleta del mundo.

Note: The same exact passage was used for conditions 1 and 2, except that for both of these conditions there appeared no instructions to focus on the specific grammatical features, and for condition 1, the relative pronouns were not highlighted in any way.

Appendix B

Examples of Pre- and Post-exposure Assessment Tasks

(1) Production of the present perfect.

Recientemente, los científicos _____ (enterarse) de que la órbita de Plutón es diferente que las de los otros planetas.
(—> se han enterado)

En los últimos días, desde el viernes pasado hasta hoy, las chicas no _____ (comer) nada.
(—> han comido)

(2) Recognition of the forms and uses of the present perfect.

Recientemente, los científicos se _____ de que la órbita de Plutón es diferente que las de los otros planetas.
a. enterarán c. enteraban
b. enteran d. han enterado
(—> d)

La profesora no va a leer la composición ahora porque ya la _____.
a. leía c. leerá
b. lea d. ha leído
(—> d)

768

(—> d)

(3) Production of the relative pronouns *que* and *quien(es)*.

Jim Thorpe era el sioux sobre _____ habló todo el mundo durante los Juegos Olímpicos de 1912.
(—> quien)

En los Estados Unidos hay muchas personas _____ están preocupadas por las condiciones ecológicas del planeta.
(—> que)

(4) Recognition of the relative pronouns *que* and *quien(es)*

Jim Thorpe era el sioux sobre _____ en los Juegos Olímpicos de 1912.
a. que habló todo el mundo
b. quien habló todo el mundo
c. él habló todo el mundo
d. aquél habló todo el mundo
(—>b)

En esta lista, se puede ver los nombres de los médicos _____.

a. cuales dan mejor servicio a sus pacientes
b. quienes dan mejor servicio a sus pacientes
c. ellos dan mejor servicio a sus pacientes
d. que dan mejor servicio a sus pacientes
(—> d)

Author

DAVID J. SHOOK, Assistant Professor, Department of Modern Languages, Georgia Institute of Technology, Atlanta, GA 30332-0375. Specializations: foreign language acquisition, teacher education, foreign language reading, grammer instruction, content language learning.

Applied Language Learning
1994, Vol. 5, No. 2, pp. 95-100

RESEARCH REPORT

Reading versus Grammar: What Students Think is Pleasurable for Language Acquisition

Jeff McQuillan
University of Southern California

Most second language (L2) acquisition theorists assign an important role to the learner's attitudes and affect in the acquisition process. This study examines the attitudes of 49 L2 students toward two language acquisition activities: grammar exercises and the extensive reading of popular literature. Students who had participated in courses based on both approaches were asked which activity was most beneficial for language acquisition, and which was more pleasurable. By a significant margin, students favored reading over grammar in both respects. Implications for L2 instructions are discussed.

Krashen (1994) has proposed a Pleasure Hypothesis for language acquisition stating that activities which are good for language acquisition are usually perceived of by acquirers as pleasant (p. 1). Other theorists and researchers have also emphasized the importance of positive affect in language acquisition (Gardner & Lambert, 1959; Mueller & Miller, 1970, Oller, Hudson, & Liu, 1977; and Schumann & Schumann, 1977). There is an abundance of evidence supporting the notion that reading is at least as effective, if not more so, than traditional grammar and skill-building activities in promoting first and second language acquisition. Krashen (1993) reviews several studies in which reading in this case, self-selected pleasure reading was compared to skill building and grammar instruction in L1 and L2 classrooms. He found that in 38 of the 41 studies, students in the reading classes did as well as or better than the comparison skill-building and/or grammar-focused groups. Elley (1991) conducted a large scale study using book-based programs in Singapore with 3,000 English as a Foreign Language (EFL) students. The experimental book-based classes outperformed students in grammer-based classes in nearly every area of language performance, including reading comprehension, grammer, listening comprehension and writing.

© 1994, Jeff McQuillan

There is also a good deal of evidence that reading is perceived as pleasurable by students who participate in L1 art programs based on literature (Sperlz, 1948; Petre, 1971; Davis & Lucas, 1971; Gray, 1969). Extensive reading, however, is not commonly used in beginning and intermediate classrooms, at least not in a foreign language (FL) setting (Huber, 1993), so there has been little opportunity to measure student preferences. This study examines how L2 students exposed to two distinct classroom activities—the learning of grammar and the extensive reading of popular literature—view them in terms of pleasure and benefits in language acquisition.

The Study

Two groups of L2 students were used in the study. Group 1 consisted of 16 third semester, university level ESL (English as a Second Language) students (M = 5, F = 11; median age = 27.5) studying at an intensive ESL program in the United States. Group 2 was comprised of thirty-three third semester, university-level SFL (Spanish as a Foreign Language) students (M = 12, F = 21; median age = 23). Both the ESL and the SFL students participated in a ten-week course in the extensive reading of popular literature in their respective second languages. All students in the study were assigned to their level on the basis of a common proficiency exam for their respective programs. Students were not informed that their section would be using popular literature as a component of the curriculum. Thus, while the selection was not random, no student was aware of the experimental nature of the courses before signing up, avoiding a self-selection bias in the sample.

Both courses consisted of the students reading newspapers, magazines, short stories, graded readers, and (for the Group 1 of ESL students) novels, and discussing them in class. The popular literature was mostly assigned, chosen by the instructors based upon previous experiences with L2 classes and student interests. Some students also read self-selected materials outside of class. Group 1 received no grammar instruction in the 10 week session; Group 2 spent less than 15 percent of class time on grammar explanations. All the students in both groups, however, had been exposed to traditional grammar and skill building activities in at least one previous or concurrent university L2 course, providing them with a basis with which to compare their literature-based course activities.

At the end of the quarter, students were asked to fill out a short survey on their preferences, indicating which activities they found pleasurable and which they found beneficial for language learning. While no pre-treatment survey was given to the two groups, previous research in other settings (Harlow & Muyskens, 1994) has found that L2 students ranked reading and grammar about the same in terms of preferred classroom activities. The present survey included close-ended questions in which students were asked to choose only one of the options listed (see questions below). This produced a count for each of the options, a count which was then subjected to statistical analysis. The three questions relevant to this study were:

(a) Based upon your experience in this class and other second language classes, which do you believe is most pleasurable: assigned readings, self-selected reading, or grammar?
(b) Based upon your experience in this class and other second language classes, which do you believe is most beneficial for your becoming fluent in your second language: assigned readings, self-selected reading, or grammar?
(c) Given a choice between reading popular literature and studying grammar, which would you prefer to do?

A Chi-square analysis of the responses of the two groups found no statistically significant differences between them, so the results were combined and are summarized in *Tables 1*, *2*, and *3*.

TABLE 1
Student Attitudes on Reading vs. Grammar: Most Beneficial (N = 49)

Assigned Popular Reading	Self-Selected Reading	Combined Reading1	Grammar Instruction
51% (25)	27% (13)	78% (38)	22% (11)

1 = Assigned Popular Reading plus Self-Selected Responses

TABLE 2
Student Attitudes on Reading vs. Grammar: Most Pleasurable (N = 49)

Assigned Popular Reading	Self-Selected Reading	Combined Reading1	Grammar Instruction
55% (27)	29% (14)	84% (41)	16% (8)

1 = Assigned Popular Reading plus Self-Selected Responses

TABLE 3
Student Preferences on Reading vs. Grammar (N = 49)

Reading	Grammar Instruction
80% (39)	20% (10)

Table 1 shows which activity students considered most beneficial. Reading is clearly the favorite, with 78 percent of the students choosing either the assigned popular reading or self-selected categories. The difference between the Combined Reading (assigned popular reading plus self selected) and Grammar Instruction categories was statistically significant from chance (Chi-square = 6.97 [Yate's Correction for Continuity applied], df = 1, $p < .01$). *Table 2* shows that students exposed to both reading and grammar also consider the former to be more pleasurable. Again, the results of the Combined Reading and Grammar cells were compared and found to differ significantly from chance (Chi-square = 11.16 [Yate's Correction for Continuity applied], df = 1, $p < .001$). Finally, *Table 3* summarizes student preferences in choosing between traditional grammar instruction and popular reading in the L2 classroom, with reading the favorite of 80 percent of the students. These results are also significantly different from chance (Chi-square = 8.23 [Yate's Correction for Continuity applied], df = 1, $p < .005$).

Discussion

The instructors for both courses noted that student enthusiasm and interest in learning the L2 was greatly improved over the general level shown by students in previous grammar-oriented courses. The preference of students for assigned versus self-selected reading may seem unusual, but the instructors reported that the vast bulk of the assigned texts were well-liked materials that had been popular with previous students. Given a choice between reading enjoyable materials supplied by the instructor and having to search for new reading matter, students may be indicating both their enjoyment of the texts supplied and their convenience. Students were not given a choice among the assigned texts, but could do additional reading which they self-selected. In any case, given an opportunity to experience both popular reading materials and grammar instruction, the overwhelming majority of the students found reading to be superior to traditional instruction not only in terms of pleasure, but also in perceived benefits in language acquisition.

It is important to point out that all of the subjects in this study had an opportunity to read popular L2 literature in and outside the classroom, a situation not common in other beginning and intermediate FL courses, as noted above. This may help explain why other measures of student preferences in L2 classroom activities differ from the present study's results. In Harlow and Muyskens (1994), for example, subjects ranked reading seventh and grammar ninth in importance in the FL classroom, suggesting only a marginal preference for reading. There is no indication, however, that students in the survey had much contact with L2 reading in their classes. Further study is needed to determine whether these preferences vary according to level. It may be, for example, that students at beginning levels perceive a greater need for grammar instruction than those in intermediate classes. Here again, however, it would be necessary to make sure students have exposure to a both reading and grammar, a point emphasized in this study.

The results lend support to Krashen's Pleasure Hypothesis, in so far as reading is both good for acquisition and perceived as more pleasurable than other equally or

less effective methods of L2 instruction, such as grammar activities. It should be noted that while Krashen states that activities which are good for acquisition are also pleasurable, the reverse is not necessarily true. An activity that is pleasurable may not be aiding in acquisition. Given the evidence in support of reading in promoting first and second language acquisition, however, it seems prudent that L2 instructors consider the use of extensive reading with popular literature in their classrooms in order to foster positive affect and more enjoyable environments for acquiring a second language.

Acknowledgment

I would like to thank Lucy Tse for her help in collecting data and for her comments on this paper. I would similarly like to thank Steve Krashen for his suggestions on an initial draft of this article.

References

Davis, F. & Lucas, J. (1971). An experiment in individualized reading. *The Reading Teacher, 24*, 737-747.

Gardner, R. & Lambert, W. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology, 13*, 266-272.

Gray, G. (1969). A survey of children's attitudes toward individualized reading. In S. Ducker (Ed.), *Individualized reading: Readings* (pp. 330-332). Metuchen, NJ: The Scarecrow Press.

Harlow, L. & Muyskens, J. (1994). Priorities for intermediate-level language instruction. *Modern Language Journal, 78*, 141-154.

Huber, B. (1993). Characteristics of college and university foreign language curricula: Findings from the MLA's 1987-89 survey. *ADFL Bulletin, 24*, 6-21.

Krashen, S. (1993). *The power of reading*. Englewood, CO: Libraries Unlimited.

Krashen, S. (1994). *The pleasure hypothesis*. Paper presented at the Georgetown Round Table on Language and Linguistics, Washington, DC.

Mueller, T. & Miller, R. (1970). A study of student attitudes and motivation in a collegiate French course using programmed language instruction. *International Review of Applied Linguistics in Language Teaching, 8*, 297-320.

Oller, J., Hudson, A., & Liu, P. (1977). Attitudes and attained proficiency in ESL: A sociolinguistic study of native speakers of Chinese in the United States. *Language Learning, 27*, 173-183.

Petre, R. (1971). Reading breaks make it in Maryland. *The Reading Teacher, 15*, 191-194.

Schumann, F. & Schumann, J. (1977). Diary of a language learner: An introspective study of second language learning. In H. D. Brown, C. Yorio, & R. Crymes (Eds.), *On TESOL '77* (pp. 241-249). Washington, DC: TESOL.

Sperlz, E. (1948). The effect of comic books on vocabulary growth and reading comprehension. *Elementary English, 25*, 109-113.

Author

JEFF MCQUILLAN, Assistant Lecturer, American Language Institute, University of Southern California, Los Angeles, CA, 90089-1294. Specializations: second language acquisition, literacy.

Calendar of Events*

1995

30 March - 2 April, *Central States Conference on the Teaching of Foreign Languages* with *Pacific Northwest Conference on Foreign Language, Southwest Conference on Language Teaching,* and *Colorado Congress of Foreign Language Teachers*, Denver. Information CSCTFL, Madison Area Technical College, 3350 Anderson Ave., Madison, WI 53704; (608) 246-6573, Fax (608)246-6880.

2 - 5 April, *Georgetown University Round Table*, Washington. Information School of Languages & Linguistics, 303 Intercultural Center, Georgetown University, Washington, DC 20057-1067.

5 - 8 April, *Languages and Communication for World Business and the Professions*, Ypsilanti. Information EMU Conference on Languages for World Business, World College, 307 Goodison Hall, Eastern Michigan University, Ypsilanti, MI 48197; (313) 487-2414, Fax (313) 485-1980.

6 - 9 April, *Association for Asian Studies*, Washington. Information Association for Asian Studies, 1 Lane Hall, University of Michigan, Ann Harbor, MI 48109.

6 - 9 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

17 - 19 April, *Exploring Language, Culture, and Literature in Language Learning*, Singapore. Information Seminar Secretariat, SEAMEO Regional Language Centre, 30 Orange Grove Road, Singapore 1025, Republic of Singapore; + (65) 737-9044, Fax + (65) 734-2753, Email [gborelc@nusvm].

28 - 29 April, *Classical Association of the Atlantic States with New Jersey Classical Association*, New Brunswick. Information Lowell Edmunds, Classical Studies Department, Rutgers University, New Brunswick, NJ 08903-0270; (908) 932-9305, Fax (908) 932-9246, Email [edmunds@elbereth.rutgers.edu].

28 - 30 April, *California Foreign Language Teachers Association*, Los Angeles. Information Hal Wingard, 6450 Lance Way, San Diego, CA 92120.

11 - 13 May, *Cincinnati Conference on Romance Languages and Literatures*, Cincinnati. Information Dept. of Romance Languages, University of Cincinnati, Mail Location 0377, Cincinnati, OH 45221-0377; (513) 556-1827 or (513) 556-1950, Fax (513) 556-2577.

12 - 14 May, *Formal Approaches to Slavic Linguistics*, Ithaca. Information Department of Modern Languages and Linguistics, Morrill Hall, Cornell University, Ithaca, NY 14853; (607) 255-0712.

20 May, *On the Work of Vercors*, Angers. Information Georges Cesborn, Centre de Recherche en Litterature et Linguistique d lAnjou et des Bocages de l'Quest, UFR Lettres, Langues et Sciences humaines, 11, Bd Lavoisier, 49045 Angers cedex; +33 41-35-21-22, Fax +33 41-35-21-91.

* Courtesy of *The Modern Language Journal* (University of Wisconsin Press)

23 - 28 May, *International Association for Learning Laboratories*, South Bend. Information IALL, University of Pittsburgh, PA 15260; (219) 631-4269, Email [IALL95@nd.edu].

1 - 3 June, *Conference on Community Interpreting*, Toronto. Information Brian Harris, School of Translators and Interpreters, Faculty of Arts, University of Ottawa, 52 rue Universit, Ottawa, Ontario, K2N 6N5, Canada.

9 - 11 June, *Teaching of Translation and Interpreting*, Elsinore. Information Cay Dollerup, Centre for Translation and Lexicography, University of Copenhagen, 96 Njalsgade, 2300 Copenhagen, Denmark.

19 - 23 June, *Computer Assisted Language Instruction Consortium*, Middlebury. Information CALICO, Duke University, 014 Language Center, Box 90267, Durham, NC 27708-0267; (919) 660-3180, Fax (919) 660-3183, Email ¦CALICO@acpub.duke.edu.

20 - 23 June, *Moscow Association of Applied Linguistics*, Moscow. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912) 333-7358, Fax (912) 333-7389, Email ¦lbradley@grits.valdosta.peachnet.edu.

26 June - 28 July, *Technology and the Human Factor in Foreign Language Education*, Honolulu. Information 1995 NFLRC Summer Institute, University of Hawai'i at Manoa, East-West Road, Building 1, Room 6-A, Honolulu, HI 96822; (808) 956-9424, Fax (808) 956-5983, Email ¦sltcc@uhunix.uhcc.hawaii.edu.

28 June - 1 July, *American Classical League Institute*, Baton Rouge. Information Geri Dutra, Americal Classical League, Miami University, Oxford, OH 45056; (513) 529-7741.

July dates to be announced, *American Association of Teachers of French*, San Antonio. Information AATF, 57 E. Armory Avenue, Champaign, IL 61820.

11 - 15 July, *Association for Computers and the Humanities with Association for Literary and Linguistic Computing*, Santa Barbara. Information Eric Dahlin, Email hcf1dahl@ucsbuxa.ucsb.edu.

12 - 14 July, *African Language Association of Southern Africa*. Information Justus C. Roux, Department of African Languages, University of Stellenbosch, Stellenbosch 7599, South Africa; +272 808-2071, Email ¦jcr@maties.sun.ac.za.

15 - 17 July, *Society for the History of Authorship, Reading and Publishing*, Edinburgh. Information Jonathan Rose, History Department, Drew University, Madison, NJ 07940.

August dates to be announced, *Computer-Supported Processing of Medieval Slavic Manuscripts*, Sofia. Information David J. Birnbaum, Department of Slavic Languages, 1421 Cathedral of Learning, University of Pittsburgh, Pittsburgh, PA 15260; Email ¦djbpitt+@pitt.edu.

4 - 7 August, *American Association of Teachers of German with Internationaler Deutschlehrerverband*, Stanford University. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398.

6 - 11 August, *International Council for Central and East European Studies*, Warsaw. Information George Fowler, Slavic Department, Indiana University, Bloomington,

IN 47405; Fax (812) 855-2107, Email !gfowler@indiana.edu.

7 - 11 August, *American Association of Teachers of Spanish and Portuguese*, San Diego. Information AATSP, Gunter Hall 106, University of Northern Colorado, Greeley, CO 80639; Fax (303) 351-1095.

7 - 9 September, *EUROCALL 95*, Valencia. Information Ana Gimeno, Departamento de Idiomas, Universidad Polit=82cnica de Valencia, Camino de Vera s/n 46071 Valencia, Spain; +34 6 387-7530 ext. 5300/5301, Fax +34 6 387-7539, Email agimeno@idm.upv.es.

10 - 12 September, *Computer Assisted Language Learning*, Exeter. Information Daphne Morton, Department of French, The University, Exeter EX4 4QH, United Kingdom; +44 19 392-264221, Fax +44 19 392-264377, Email k.c.cameron@exeter.ac.uk.

15 - 16 September, *Reading in the University in First, Second and Foreign Languages*, Toulouse. Information Gail Taillefer, Department des Langues et Civilisations, Universit des Sciences Sociales, Place Anatole-France, 31042 Toulouse, France; Fax +33 61-22-94-08.

21 - 24 September, *German Studies Association*, Chicago. Information Helga A. Welsh, Department of Politics, Reynolda Station, Box 7568, Wake Forest University, Winston-Salem, NC 27109; (910) 759-5452, Fax (910) 759-6104, Email welsh@wfu.edu.

October dates to be announced, *Iowa Foreign Language Association*, Des Moines. Information Paul Hoekstra, Bureau of Curriculum & Instruction, Grimes State Office Building, Des Moines, IA 50319.

5 - 7 October, *Purdue University Conference on Romance Languages, Literatures and Film*, West Lafayette. Information Department of Foreign Languages, 1359 Stanley Coulter Hall, Purdue University, West Lafayette, IN 47907-1359; (317) 494-7691, Fax (317) 496-1700, Email rla@vm.cc.purdue.edu.

5 - 7 October, *Mountain Interstate Foreign Language Conference*, Radford. Information Leonor Ulloa, Department of Foreign Languages and Literatures, P.O.Box 6937, Radford University, Radford, VA 24142; (703) 831-5120, Fax (703) 831-5970, Email lulloa@ruacad.ac.runet.edu.

6 - 7 October, *Pennsylvania Modern Language Association*, Pittsburgh. Information John Pearce, Muhlenberg College, Foreign Languages and Literatures, Allentown, PA 18104-5586.

6 - 7 October, *Classical Association of the Atlantic States with New Jersey Classical Association*, Atlantic City. Information John C. Traupman, 201 Tower Lane, Penn Valley, PA 19072-1127; (610) 664-5487.

6 - 9 October, *Second Language Research Forum*, Ithaca. Information Nuria Lopez Ortega, Cornell University, Department of Modern Languages and Linguistics, Morrill Hall, Ithaca, NY 14853-4701; Email slrf95@cornell.edu.

7 - 8 October, *Carolinas Symposium on British Studies*, Boone. Information Jacqueline Gmuca, Department of English, Coastal Carolina University, Conway, SC 29526.

12 - 14 October, *Penn State Conference on the Acquisition of Spanish as a First or Second Language*, State College. Information William R. Glass, Pennsylvania State University, Department of Spanish, Italian, and Portuguese, 352 N. Burrowes Building, University Park, PA 16802; Email wrg3@psuvm.psu.edu.

12 - 14 October, *Foreign Language Association of North Ca*
mation Wayne Figart, 204 N. 16th St., Wilmington, NC 2
13 - 14 October, *Washington Association for Foreign Lang*
Information Rosemary Leiva, 1412 S. 334d Ave., Yakima,
or (509) 965-2040.
18 - 20 October, *Foreign Language Association of North D*
Herbert Boswau, PO box 8198, Grand Forks, ND 58202-8
19 - 20 October, *Minnesota Council on the Teaching of Fore*
be announced. Information Millie Mellgren, 3055 Rose
55441.
19 - 21 October, *Classical Association of the Empire State*,
Crooker, 15 Sutherland Street, Pittsford, NY 14534.
20 - 22 October, *New York State Association of Foreign Lang*
Lake. Information Helene Combopiano, 71 Grand Bo
13905.
26 - 28 October, *Texas Foreign Language Association*, Aus
Simons, 1320 Modiste Drive, Houston, TX 77055.
26 - 28 October, *South Central Modern Language Associat*
Executive Director, SCMLA, Texas A&M University, C
4227.
26 - 29 October, *American Association for the Advancemen*
ington. Information AAASS, Stanford University, Jord
Stanford, CA 94305-4130; (415) 723-9668.
2 - 5 November, *Japan Association of Language Teachers*,
Information JALT, Glorious Tokyo #301, 2-32-10 Nishi
116, Japan; +81 3 3802-7121, Fax +81 3 3802-7122.
3 - 4 November, *National Association of Self-Instructional La*
ington. Information NASILP, Critical Languages, 022-38
delphia, PA 19122.
18 - 20 November, *American Council on the Teaching of*
American Association of Teachers of German, Anahei
Executive Plaza, Yonkers, NY 10701-6801; (914) 963-883
30 November - 2 December, *Formal Description of Slavic Lan*
tion Gerhild Zybatow, Universität Leipzig, Philologis
Slavistik, Augustusplatz 9, 04109 Leipzig, Germany; +034
719 3002, Email slavlips@rzaix340.rz.uni-leipzig.de.
5 - 8 December, *Conference on Computers in Education*, Sin
95/AACE, P.O. Box 2966, Charlottesville, VA 22902; (804
7449, Email AACE@virginia.edu.
27 - 30 December, *American Association of Teachers of Sl*
guages & Literatures, Chicago. Information AATSEE
guages, University of Arizona ML 340, Tucson, AZ 857
27 - 30 December, *Modern Language Association of Americ*

ion of North Carolina, Greensboro. Infor-
lmington, NC 28401; (919) 763-4009.
r Foreign Language Teaching, Portland.
d Ave., Yakima, WA 98902; (509) 457-0966

tion of North Dakota, Fargo. Information
rks, ND 58202-8198.
eaching of Foreign Languages, location to
gren, 3055 Rosewood Lane, Plymouth, MN

Empire State, Syracuse. Information Jill
NY 14534.
of Foreign Language Teachers, Kiamesha
o, 71 Grand Boulevard, Binghamton, NY

ssociation, Austin. Information Eugenia
TX 77055.
guage Association, Houston. Information
M University, College Station, TX 77843-

he Advancement of Slavic Studies, Wash-
University, Jordan Quad/Acacia Building,
8.
uage Teachers, location to be announced.
, 2-32-10 Nishinippori, Arakawa-ku Tokyo
802-7122.
-Instructional Language Programs, Wash-
nguages, 022-38, Temple University, Phila-

he Teaching of Foreign Languages with
erman, Anaheim. Information ACTFL, 6
1; (914) 963-8830, Fax (914) 963-1275.
ion of Slavic Languages, Leipzig. Informa-
pzig, Philologische Fakultät, Institut für
Germany; +0341 719 2499/3003, Fax +0341
-leipzig.de.
Education, Singapore. Information ICCE
, VA 22902; (804) 973-3987, Fax (804) 978-

f Teachers of Slavic & E. European Lan-
nation AATSEEL, Russian & Slavic Lan-
Tucson, AZ 85721.
iation of America, Chicago. Information

MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863.

1996

29 February - 2 March, *Southern Conference on Language Teaching with Alabama Association of Foreign Language Teachers*, Mobile. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912) 333-7358, Fax (912) 333-7389, Email ¦lbradley@grits.valdosta.peachnet.edu¦.

5 - 8 March, *National Association of Independent Schools*, Washington. Information Claire Whalen, NAIS, 1620 L Street NW, Washington, DC 20036.

26 - 30 March, *Teachers of English to Speakers of Other Languages*, Chicago. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2705.

28 - 31 March, *Central States Conference on the Teaching of Foreign Languages*, Louisville. Information CSCTFL, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; (608) 246-6573, Fax (608) 246-6880.

11 - 13 April, *Southwest Conference on Language Teaching*, Albuquerque. Information Joann Pompa, Mountain Pointe High School, 4201 E. Knox Road, Phoenix, AZ 85044; (602) 759-8449 Ext. 3036, Fax (602) 759-8743.

11 - 14 April, *Association for Asian Studies*, Honolulu. Information Association for Asian Studies, 1 Lane Hall, University of Michigan, Ann Arbor, MI 48109.

18 - 21 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

26 - 27 April, *Classical Association of the Atlantic States*, Baltimore. Information Sister Therese Marie Dougherty, Notre Dame College of Baltimore, 4701 N. Charles Street, Baltimore, MD 21210; (301) 435-0100.

26 - 28 April, *California Foreign Language Teachers Association*, Sacramento. Information Hal Wingard, 6450 Lance Way, San Diego, CA 92120.

July dates to be announced, *American Association of Teachers of French*, France. Information AATF, 57 E. Armory Avenue, Champaign, IL 61820.

4 - 9 August, *International Association of Applied Linguistics*, Jyv skyl Finland. Information Udo O.H. Jung, Language Centre, University of Bayreuth, D-95440 Bayreuth, Germany; (0921) 553595, Fax (0921) 553594, Email ¦udo.jung@uni-bayreuth.de¦.

29 - 31 August, *EUROCALL 96*, Szombathely, Hungary. Information June Thompson, CTI Centre for Modern Languages, University of Hull, Hull HU6 7RX, United Kingdom; Email ¦Eurocall@hull.ac.uk¦.

17 - 18 October, *Minnesota Council on the Teaching of Foreign Languages*, location to be announced. Information Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

18 - 20 October, *New York State Association of Foreign Language Teachers*, Kiamesha Lake. Information Helene Combopiano, 71 Grand Boulevard, Binghamton, NY 13905.

18 - 20 October, *Texas Foreign Language Association*, Houston. Information Eugenia Simons, 1320 Modiste Drive, Houston, TX 77055.

23 - 25 October, *Foreign Language Association of North Dakota*, Bismarck. Information Herbert Boswau, PO box 8198, Grand Forks, ND 58202-8198.

24 - 26 October, *Foreign Language Association of North Carolina*, Greensboro. Information Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; (919) 763-4009.

1 - 2 November, *National Association of Self-Instructional Language Programs*, Washington. Information NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19122.

22 - 24 November, *American Council on the Teaching of Foreign Languages with American Association of Teachers of German*, Philadelphia. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275.

27 - 30 December, *Modern Language Association of America*, location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863.

1997

3 - 6 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

10 -12 October, *Foreign Language Association of North Carolina*, Greensboro. Information Wayne Figart, 204 N. 16th St., Wilmington, NC 28401; (919) 763-4009.

10 - 13 April, *Central States Conference on the Teaching of Foreign Languages*, Columbus. Information CSCTFL, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; (608) 246-6573, Fax (608) 246-6880.

5 - 7 August, *Problems of Teaching Modern Hebrew*, Jerusalem. Information Ben-Zion Fischler, Council on the Teaching of Hebrew, P.O.B. 7413, Jerusalem 91073, Israel.

27 - 30 December, *Modern Language Association of America*, location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863.

1998

September dates to be announced, *International Congress of Slavists*, Cracow. Information Robert Rothstein, Department of Slavic Languages, Herter Hall, University of Massachusetts, Amherst, MA 01003-3940.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (ALL) is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. ALL publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editor encourages the submission of original manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assesment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784.
Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

The list of references should be submitted on a separate page of the manuscript with the centered heading: *References*. The entries should be arranged alphabetically by surname of authors. The sample list of references below illustrates format for bibliographic entries:

Dulay, H. & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly*, *16* (1), 93-95.

Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

Reference citations in the text of the manuscript should include the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that is being quoted originally appeared, e.g., (Jones, 1982, pp. 235-238). All works cited in the manuscripts must appear in the list of references, and conversely, all works included in the list of references must be cited in the manuscript.

Notes should be used for substantive information only, and they should be numbered serially throughout the manuscript. Subsequently they all should be listed on a separate page title *Notes*.

Specifications for Floppy Disks

Where feasible, manuscripts are preferred on 3.5" disk. Manuscripts produced on one of the DOS-based or Macintosh systems should be formatted as MS-DOS file at double density, if possible.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____

2. Disk is formatted as: double__ high density__

3. Remarks: _____

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the sub_ ct of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

784

PIN: 072639-1/N-000

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

786

Professional Bulletin 65-94-2
United States Army
Approved for public release.
Distribution is unlimited.

786

Applied Language Learning

VOLUME 6 • NUMBER 1&2
1995

787

Applied Language Learning

Volume 6 • Numbers 1 & 2

1995

Defense Language Institute
Foreign Language Center
Presidio of Monterey

788

Applied Language Learning
PB 65-95-1 & 2

The mission of Professional Bulletin 65-95, *Applied Language Learning* (US ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semi-annually by the Defense Language Institute, Foreign Language Center and Presidio of Monterey, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

DENNIS J. REIMER
General, United States Army
Chief of Staff

Official:

JOEL B. HUDSON
Acting Administrative Assistant to the Secretary of the Army
00732

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to DLIFLC. *Applied Language Learning* is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster:

Send change-of-address information to:

Applied Language Learning
Office of the Provost
Defense Language Institute, Foreign Language Center, and Presidio of Monterey
Monterey, CA 93944-5006
Telephone: 408 242-5638
DSN: 878-5638

Applied Language Learning

PB-65-95-1&2
USSN 1041-6790

Lidia Woytak
Editor

Editorial Board

Kathleen M. Bailey	Professor, MIIS
David P. Benseler	Professor, Case Western Reserve University
Christine Campbell	Project Director, DLI
John L. D. Clark	Dean, DLI
Ray T. Clifford	Provost, DLI
Martha Herzog	Asssociate Provost, DLI
Stephen D. Krashen	Professor, USC
John A. Lett, Jr.	Director, Research, DLI
June K. Phillips	Dean, Weber State University
Leo van Lier	Professor, MIIS

DLI Staff

Editorial Coordinator, Christa Rutsche
Associate Editor, Steven Soloman
Software Consultants, Milan Herben, John Jaworski
Graphics, Steven Jackson
Wordprocessing & Distribution, Bonnie Matheney

790

A Note of Appreciation

The individuals listed below served as referees of manuscripts submitted to *Applied Language Learning* in 1995. The editor and the editorial board express their gratitude for expert service to:

Kathleen M. Bailey
Martha Bean
Christine Campbell
Patricia L. Carrell
John L. D. Clark
Robert C. Gardner
John S. Hedgecock
Martha Herzog
Eli Hinkel
Gordon Jackson
Gisèle Kapuscinski
Gabriele Kasper
Steven Krashen
B. Kumaravadivelu
James P. Lantolf
James F. Lee
John A. Lett, Jr.
Leo Van Lier
Martha Nyikos
Rebecca L. Oxford
Howard Roland
Richard Schmidt
Richard G. Tucker
Dolly J. Young

Applied Language Learning

1995 Volume 6 Numbers 1 & 2

Articles

1 Examining the Value of Conversation Partner Programs
Fredricka L. Stoller, Rosemary Hodges, and *Jeania Kimbrough*

13 Discontinuous Interlanguage Development and Its Implications for Oral Proficiency Rating Scales
Richard Young

27 Silence, Interruptions, and Discourse Domains: The Opportunities to Speak
Raimundo Mora

41 Pragmatics in Interlanguage: German Modal Particles
Martina Möllering and David Nunan

65 Sociolinguistic Patterns of Arabic Native Speakers: Implications for Teaching Arabic as a Foreign Language
Anwar S. Hussein

Reviews

89 Cook: Discourse.....*M.S. Bean*

91 Van Lier: Introducing Language Awareness.....*L. Woytak*

General Information

95 Calendar of Events

99 Information for Contributors

Applied Language Learning
1995, Vol. 6, Nos. 1 & 2, pp. 1-12

Examining the Value of Conversation Partner Programs

Fredricka L. Stoller
Rosemary Hodges
and
Jeania Kimbrough
Northern Arizona University

The survey study reported here was designed to examine the value of conversation partner programs which pair up second language students and native speakers for out-of-class conversation. Data analyses shed light on activities conversation partners engage in, the perceived advantages of the conversation partner experience for nonnative and native participants, and the problems encountered by conversants. Results can be used to help language programs plan, organize, and market their conversation partner programs better; in addition, insights gained can be used to ensure more successful partnerships by incorporating survey results into written guidelines and/or introductory orientation sessions for conversation partner program participants.

Many language programs organize and administer conversation partner programs which pair up second language (L2) students and native speakers for out-of-class conversation. They do so because, at the intuitive level, teachers and administrators believe that conversation partner programs have multiple benefits for L2 students. The perceived advantages include opportunities for L2 students to develop: (a) improved language proficiency, (b) refined conversational strategies, (c) an enhanced understanding of and accelerated adjustment to the target culture, and (d) a secure social outlet (Duryee, Lanier, & Michel-Reyes, 1991; Fragiadakis & Licwinko, 1986; Higgins, 1992; Schmidt & Frota, 1986; Tacoma Community House, 1985). The casual and friendly interchanges that occur also lead to judgment-free acceptance of the language learner as a whole human being, resulting in improved self esteem for the nonnative speaker (J. Zukowski-Faust, personal communication, November 1994).

Supporting these intuitions are studies which illustrate how interaction between native speakers and nonnative speakers facilitates learner comprehension, access to lexical form and meaning, production of modified output, hypothesis testing,

© 1995, Fredricka L. Stoller, Rosemary Hodges, and Jeania Kimbrough

and development of strategies to overcome miscommunication (see Brock, Crookes, Day, & Long, 1986; Duryee, Lanier, & Michel-Reyes, 1991; Pica, 1994). The results of these studies suggest that conversation partner programs can help nonnative speakers improve both their conversational strategies and language proficiency. (See Pica, 1994, for a comprehensive review of a decade of research on social interaction and negotiation of L2 learners and their interlocutors.)

Other studies, however, question the value of informal conversation for adult language learners. Duryee, Lanier, and Michel-Reyes (1991) cite claims that poorly trained (or untrained) native speakers have trouble modifying their speech for low proficiency L2 learners, causing either miscommunication or no communication at all. Despite these claims, Higgins (1992), in a study of communication as social action, concludes that conversation partners do, in fact, modify their output when speaking with nonnative speakers to achieve a "shared reality" with them. Freed (1980) found that adult native speakers show regard for foreign peers as cognitive equals even though their responses to foreigners and children were similar syntactically.

Native/nonnative speaker interaction does not guarantee fluency or idiomaticity, "just as instruction does not guarantee grammaticality" (Schmidt & Frota, 1986, p. 307). The fact remains, however, that nonnative speakers themselves believe that having conversational interaction with native speakers enhances their L2 development (Duryee, Lanier, & Michel-Reyes, 1991). Consequently, many L2 learners are eager to participate in conversation partner programs that facilitate native and nonnative speaker interaction. Despite the potential problems associated with organized partnerships (e.g., lack of time due to busy schedules, mismatch of partners, unbalanced use of time during conversational exchange), having a conversation partner program in an L2 curriculum is attractive to L2 learners.

Conversation Partner Program Formats

A review of the ESL conversation partner program literature reveals six basic conversation partner program formats[1]:

1. Nonnative speakers of English are paired with native speakers of English who are studying a foreign language (specifically the first language of their conversation partner). In this type of program, partners devote equal time to both languages so that each partner has the opportunity to practice and improve his/her L2 skills (Fragiadakis & Licwinko, 1986).

2. Nonnative speakers of English are matched with individuals from the campus community: staff, students studying a foreign language, and/or graduate students in MA-TESL or related programs (Duryee, Lanier, & Michel-Reyes, 1991).

3. Nonnative and native speakers of English are introduced through a cross-cultural communication course (or other college/university course); both participants receive credit for the course, though the course may or may not be mandatory for the native speakers (Duryee, Lanier, & Michel-Reyes, 1991).

4. Nonnative speakers of English, often new immigrants or refugees, are in-

troduced to community volunteers who may or may not be connected with nearby educational institutions (Tacoma Community House, 1985).

5. Nonnative speakers of English are paired up with conversants who are paid by the language program administering the conversation partner program (Kimbrough & Hodges, 1993).

6. Nonnative speakers are paired with senior citizens for a telephone-partner program (see Stoller, 1992). Initial contact is established by letter and is followed up by weekly telephone calls.

The Present Study

The survey study reported here was designed to examine the value of one conversation partner program format, specifically the type described in the second category above; that is, in this study respondents were: (a) international students enrolled in an intensive English program and (b) graduate students in an MA-TESL program. The primary impetus for the study was to gather information that could be used to: (a) enhance written guidelines and orientation sessions for conversation program participants, (b) guide participants towards more successful partnerships, and (c) steer participants away from potential pitfalls. In addition, the study was conducted with the hope that results would prove beneficial for educators and program administrators involved with other conversation partner program formats as well.

Survey respondents were asked to answer four questions. Although few in number, the questions targeted information that was considered most essential to understanding the keys to successful partnerships. The brevity of the questionnaire served two other purposes as well: Participants were not put off by a lengthy survey and nonnative respondents were not intimidated by the language required to successfully complete the survey. The questions were as follows:

1. What types of activities do you enjoy doing most with your conversation partner?
2. In what ways did *you* benefit from the Conversation Partner Program?
3. In what ways do you think *your partner* benefited from the Conversation Partner Program?
4. What problems, if any, did you and/or your partner experience?

Method

Subjects

Subjects for this study were international students enrolled in the Program in Intensive English (PIE) at Northern Arizona University (NAU) and native (and near-native) speakers of English who were enrolled as graduate students in the MA-TESL program at NAU. All subjects (34 PIE students and 34 MA-TESL students) had volun-

teered to be participants in the Conversation Partner Program and had agreed to meet at least one hour per week during a 15-week semester.[2]

Procedures

The four-question survey was distributed to native and nonnative students who participated in the Conversation Partner Program during the 1993-1994 academic year and the Fall 1994 semester. Sixty-eight surveys were distributed; sixty-three were returned, resulting in a 93% return rate. Thirty-four of the completed surveys were from nonnative speaker (NNS) participants and twenty-nine were from native speaker (NS) participants.

The survey was designed to generate open-ended responses; respondents could provide as many comments as they wanted for each survey item. As a result, the total number of responses for each question often exceeded the number of respondents in the study. Initially two of the three researchers, on an individual basis, grouped responses for each question into what each researcher perceived to be non-overlapping categories. Discrepancies were discussed by all three researchers until they agreed on the categorization of responses. After discussing and analyzing each cluster of responses, the research team assigned a "label" to each category, one which captured the essence of the group of responses as a whole.

Results

Data analyses shed light on activities conversation partners engage in, the perceived advantages of the conversation partner experience for both nonnative and native participants, and the problems encountered by conversation partners. Results are summarized below.

Activities

Survey results reveal the settings in which conversational exchange takes place and the specific activities in which participants engage.[3] Survey respondents indicated that having coffee and eating meals together were the most common and enjoyable conversation partner activities because partners could easily discuss current events, controversial topics, and personal issues in such settings. Figure 1 lists other activities of the respondents during their semester-long partnerships.

Benefits for Nonnative Speakers

As indicated in Table 1, an analysis of survey responses—from both native and nonnative participants—reveals that the perceived benefits of conversation partner program participation for *nonnative speakers* can be divided into three general areas.

Evening Activities

go out to dinner
fix dinner for each other
go to movies/watch videos
go to bars and clubs (dancing)
host or attend parties
go to concerts and music festivals
go to local observatory, fireworks display, etc.
play games (e.g., hearts, twister, charades)

Daytime Activities

go out for coffee, breakfast, lunch
go shopping (e.g., clothing, groceries)
run errands (e.g., hairdresser, travel agency)
go bowling, skating, weight-training
study/do homework
visit radio stations, newspaper offices, etc.

Outdoor Activities

go for hikes and walks
participate in sports (e.g., basketball, skiing)
play games (e.g., hackysack, frisbee)
take day trips
go camping
have cookouts, barbecues, and picnics
join outdoor recreational clubs together
(e.g., university Hiking Club, Ski Club)
go horseback riding

Cultural Activities

visit each other's homes
attend lectures, workshops, club meetings
attend plays, symphonies, festivals, carnivals
visit museums, galleries, exhibitions
visit each other's classes
attend church

Holiday-Related Activities

attend holiday parties
bake holiday treats
celebrate holidays together (by carving pumpkins,
preparing Thanksgiving dinner, exchanging
Christmas presents, decorating a Christmas tree,
teaching/singing holiday songs, making holiday arts and crafts)

FIGURE 1. *Activities conversation partners participate in.*

Note. Most frequently noted activities are listed first in each category. Others are listed in descending order of frequency.

First, conversation partner programs provide a social and emotional outlet for nonnative participants; second, they lead to language improvement; and third, they build cultural sensitivity. The number of responses in each one of the categories allows for the more detailed analyses provided below.

TABLE 1
Perceived benefits of conversation partner program (CPP) for nonnative speakers (NNS)

Benefits of CPP for NNS	As perceived by NNS*	As perceived by NS*
Provides a social/emotional outlet	27	22
Leads to language improvement	25	16
Builds cultural sensitivity	18	26

*Numbers correspond to number of responses rather than number of respondents.

Social and emotional outlet. As shown in Table 1, nonnative respondents perceive the social and emotional outlet provided by the Conversation Partner Program to be its greatest advantage. The most frequently cited social/emotional benefit for nonnative participants was the opportunity to meet someone to go out with and have fun with. The opportunity to make "an instant friend" was cited as another valuable benefit. As one nonnative speaker put it, "The conversation partner could be a good friend who can share happiness and depressing with." Similarly, one native participant stated that through the conversation Partner Program her nonnative partner gained a "confidante, a shoulder to cry on when she was having problems with her host family." A nonnative speaker said, "I enjoyed talking about our lives, our families, and our cultures with a cup of coffee."

Although not identified as an advantage by nonnative speakers themselves, native speakers noted that one benefit of the Conversation Partner Program was that it gave the nonnative speaker an opportunity to speak with someone about topics which are taboo in L1 cultures. For instance, one native speaker stated that because her "conversation partner had an American boyfriend, she was very interested in knowing about American perspectives on [male/female] relationships. She told me she couldn't talk about her relationship [with an American male] with her Japanese girlfriends."

Language improvement. Nonnative participants appreciate the opportunities for out-of-class conversation that result from Conversation Partner Program participation. One international student expressed it this way: "The Conversation Partner Program is an ideal way to acquire the English language because we learn from a [native]

speaker rather than learning it in English class." This comment reflects the intuitive sense that conversation partner programs help L2 learners improve their conversational skills. In fact, many respondents noted that the speaking and listening opportunities that resulted from partner interchanges led to improvement in the following areas:

1. fluency
2. accuracy in pronunciation and word choice
3. vocabulary development
4. confidence
5. sociolinguistic and strategic competence

Although the nonnative participants did not use the term "strategic competence" in their survey responses, the concept was evident when one nonnative participant identified the need to talk louder when conversing with her conversation partner and other native English speakers: "I learned that I have to speak up when I talk to English speakers. Otherwise, we cannot make friendship each other."

Building cultural sensitivity. Building cultural sensitivity was identified as another benefit of conversation partner programs for nonnative participants. The majority of nonnative respondents noted that they gained an enhanced awareness of the target culture because of the Conversation Partner Program. "It is very useful to know American culture and understand their thoughts," wrote one international student. Another stated that he had "learned American ways and different ideas towards everyday life and people."

Another respondent claimed that the Conversation Partner Program was "not only a means of learning language but also a means of exchanging cultural differences and ideas," highlighting the value of cross-cultural exchanges and exposure to different perspectives. It was noted by a number of native speakers that the conversation partner experience eases nonnative speaker adjustment into the target culture as well as yields experiences that nonnative speakers cannot attain within their own social circles. One nonnative participant, commenting on adjustment to the L2 culture, put it this way: "I learned about American culture and how to adjust with American society." One native speaker commented that her nonnative partner gained an appreciation for "Black English and Black people's culture" as a result of their partnership.

Benefits for Native Speakers

Survey responses from both native and nonnative speakers show that perceived benefits for *native speakers* fall into four general categories: social outlet, enhanced cultural sensitivity, an ability to connect English language teaching theory to practice, and exposure and resultant sensitivity to L2 learner needs. (See Table 2.) The first two categories reveal benefits that are parallel to those listed for nonnative participants but the last two disclose quite different benefits. Each of the four major catego-

ries is analyzed in greater detail below.

TABLE 2
Perceived benefits of conversation partner program (CPP) for native speakers (NS)

Benefits of CPP for NS	As perceived by NNS*	As perceived by NS*
Provides a social outlet	8	33
Builds cultural sensitivity	27	24
Connects ELT theory to practice	5	11
Exposes and sensitizes NS to L2 learner needs	6	7

*Numbers correspond to number of responses rather than number of respondents.

Social outlet. Judging by survey responses, most native speakers highly value the social aspects of the Conversation Partner Program. Because the native-speaking participants were graduate students, they regarded the Conversation Partner Program as a "nice digression from academic life." Many of the partnerships were successful in fostering friendships that would continue after the Conversation Partner Program ended for the semester. One participant stated that she "met one of the nicest people on the planet and a friend for life."

A number of respondents felt that the Conversation Partner Program provided participants with an opportunity to enhance interpersonal skills. One native speaker expressed it this way: "I had to act as host to someone else [and] try to help a foreigner, which was good for me." Similarly, two native speakers mentioned that having the opportunity to meet *other* international students through their conversation partners was particularly valuable not only in relation to social needs, but also because as future ESL/EFL teachers, they could benefit from increased exposure to L2 learners.

Enhanced cultural sensitivity. Many respondents claimed that the Conversation Partner Program enhanced native participants' cultural sensitivity by providing natural opportunities to (a) learn about another culture, (b) share culture-specific experiences, and (c) be exposed to diverse perspectives on life. One native speaker wrote that the conversant experience provided her with "an excellent opportunity to discover the similarities and differences between two countries" and also to develop a friendship. One nonnative participant stated that he shared a lot of information about his

890

culture—including facts about his language, culture, and country's history—with his partner.

Connection of theory to practice. The native speaker participants appreciated the chance to apply what they were learning in their MA-TESL classes to their conversation partner encounters. Five respondents mentioned that the Conversation Partner Program provided means of practicing their "craft"; one native speaker enjoyed having the opportunity to be "creative in teaching [her partner] English." Several native speakers appreciated being able to engage their nonnative partners in language acquisition research projects. "On a practical level, [my partner] was useful for my ESL Methodology needs analysis assignment. But also I enjoyed...spending time with [my partner] because she was eager, happy, and fun to talk to."

A number of native speakers remarked that they valued the experience of observing their partner's language development. One respondent claimed that it was "rewarding to observe her [partner's] language development—in terms of vocabulary and some idiomatic expressions—and her personal growth in the U.S."

Exposure and resultant sensitivity to L2 learner needs. The opportunity to interact with L2 learners provided a strong incentive for native speakers to participate in the Conversation Partner Program. Three native speakers noted that their participation in the Conversation Partner Program enhanced their understanding of nonnative speakers' linguistic and social needs. It is interesting to note that nonnative speakers felt that native speakers attained a greater level of comfort with nonnative speakers through participation in the Program. One native speaker pointed out that becoming familiar and comfortable with nonnative speakers' accents was of great benefit to her.

Problems

Both native and nonnative speakers indicated that they encountered some problems while participating in the Conversation Partner Program. The most frequently noted problem was a lack of time to meet with conversation partners. Many participants found it difficult to work around busy (and sometimes conflicting) class schedules. One native speaker remarked that her conversation partner's "continuous studying got in the way of a successful partner interchange."

As might be expected, several nonnative speakers mentioned feeling uncomfortable and insecure because of their "weak" communication skills. "My English is too poor to communicate smoothly...I usually misunderstood the contents but I sometimes pretended understanding [my partner's] English." Three participants referred to specific instances involving miscommunication, noting problems with "supposedly" agreed upon meeting times and places. "At first, sometimes I misunderstood when/where I should meet [my partner] because it was difficult for me to talk by phone."

Another problem encountered by a few participants was linked to a perceived lack of activities to engage in and topics to discuss. "We sometimes ran out of things

to say—especially in the beginning. I've never been a big talker so it was sometimes difficult for me to carry the conversation through lulls." One nonnative participant commented that "sometimes we couldn't find good activities or topics which we talk about."

The perception of indifferent attitudes towards one's partner was noted as a problem by one nonnative speaker and three native speakers. One nonnative speaker said that her "conversation partner might have thought that I did not need him because I already had been in U.S. for two and a half years. Thus, he had negative attitude to do some activities with me." One native speaker had this to say: "I think that [my partner] is a bit shy, therefore, I needed to prompt the meetings and conversations. I was not always certain that he wanted to meet and practice English."

Other problems mentioned by individual participants included forgetting about appointments, feeling guilty about using nonnative partners as informants for class projects, and the mismatch of personalities.

Conclusion

The benefits of the conversation partner program format examined in this study are many and varied. In addition to the language improvement often associated with conversation partner interactions, nonnative participants benefit from the social and emotional outlet that results from conversant exchanges. In addition, nonnative participants develop an enhanced cultural sensitivity to the target culture through conversation partnerships. Native speakers also benefit from the social opportunities afforded by conversational exchange and the cultural sensitivity that comes with exposure to new and different perspectives. Furthermore, the native speakers who participated in this study were able to enhance their graduate studies through conversation partner program participation because the interactions allowed them to connect theory to practice and observe L2 learners in the process of mastering a second language. The combination of social benefits and opportunities for professional growth that accompanied participation for these native speaker volunteers may have influenced their enthusiasm for the program.

It is suggested that programs similar to the one reported here use these survey findings to plan, organize, and "market" their conversation partner programs more effectively. To ensure successful partnerships, survey results can be incorporated easily into promotional materials, written guidelines, and/or orientation sessions. By providing participants with lists of activities they might want to engage in, topics they might want to discuss, and advice for avoiding possible pitfalls, participants are more likely to have positive conversation partner program experiences, ones that will lead to the multiple benefits outlined earlier in this article.

The study reported here has implications for other conversation partner program types as well, though further study is needed to determine the ways in which the other program formats are similar to and different from one another. It is suggested that administrators of other types of conversation partner programs distribute a similar

survey to native and nonnative participants in order to gather parallel data on their programs. Such studies would allow us to identify common features that lead to positive conversational exchange outcomes and to isolate characteristics that deter from the success of conversation partner programs.

Acknowledgments

The authors would like to thank William Grabe and Jean Zukowski-Faust of Northern Arizona University for their helpful feedback on earlier versions of this paper. We would also like to thank two anonymous reviewers for their comments.

Notes

[1]Similar formats exist for conversation partner programs that pair up foreign language students with native speakers of those languages.

[2]Partners decided how they would spend their time and what topics to discuss on their own.

[3]The locale of a conversation partner program and the participants themselves will influence the types of activities conversants engage in. Higgins (1992) points out that topics of conversation between partners are influenced by partners' likes and dislikes which contribute to a shared sense of social and cultural reality. Duryee, Lanier, and Michel-Reyes (1991) note that topics of conversation often focus on daily events such as holiday traditions, national time changes, restaurant etiquette, dating practices, etc. At times, topics are determined by teachers who wish to coordinate classroom activities with conversation time (Tacoma Community House, 1985).

References

Brock, C., Crookes, G., Day, R., & Long, M. (1986). Differential effects of corrective feedback in native speaker-nonnative speaker conversation. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition,* 229-236. Rowley, MA: Newbury House.

Duryee, P.P., Lanier, L. K., & Michel-Reyes, R. (1991). Conversation partners—Useful language input? An informal summary of research. (ERIC Document #ED 338 039).

Fragiadakis, H., & Licwinko, A. (1986). The conversation exchange program: Foreign and American students together. (ERIC Document #ED 274 194).

Freed, B. (1980). Talking to foreigners vs. talking to children: Similarities and differences. In R. C. Scarcella & S. D. Krashen (Eds.), *Research in second language acquisition: Selected papers of the Los Angeles SLA Forum,* 19-27. Rowley, MA: Newbury House.

Higgins, E. T. (1992). Achieving 'shared reality' in the communication game: A social action that creates meaning. *Journal of Language and Social Psychology,*

11(3), 107-131.

Kimbrough, J., & Hodges, R. (1993). Models, activities, and benefits for non-native/native speaker conversation partner programs. Poster session at NAFSA Region II Conference, Scottsdale, AZ.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44* (3), 493-527.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition,* 237-326. Rowley, MA: Newbury House.

Stoller, F. L. (1992). Taxonomy of intensive English program innovations. *Journal of Intensive English Studies, 6*, 1-25.

Tacoma Community House, WA. (1985). Talk time coordinator handbook. (ERIC Document #ED 316 024).

Authors

FREDRICKA L. STOLLER, Assistant Professor, English Department, and Director of the Program in Intensive English at Northern Arizona University, PO Box 6032, Flagstaff, AZ 86011-6032 USA. Specializations: L2 curriculum and methodology; L2 teacher education; content based instruction; innovation in ELT settings.

ROSEMARY HODGES, Soros Foundation Fellow in Moldova. Previous teaching experience in Northern Arizona University's Program in Intensive English. Specialization: English as a foreign language instruction.

JEANIA KIMBROUGH, Soros Foundation Fellow in Slovakia. Previous teaching experience in Northern Arizona University's Program in Intensive English. Specialization: English as a foreign language instruction.

Applied Language Learning
1995, Vol. 6, Nos.1 & 2, pp. 13-26

Discontinuous Interlanguage Development And Its Implications for Oral Proficiency Rating Scales

Richard Young
University of Wisconsin-Madison

Recent research in educational measurement and in second language acquisition suggests that a learner's competence in a second language may best be viewed as multidimensional. Related research has also shown that, in many respects, development of competence in a second language is not a linear process, but instead shows patterns of discontinuity and U-shaped behavior. Such a view of second language competence has important implications for language testing and, in particular, for the development of rating scales for assessing learners' competence in oral interviews. This paper critically examines two rating scales for the assessment of oral proficiency. Both scales are judged to be inadequate because they are based on an untenable theory of the development of proficiency in a second language and on mistaken intuitions about oral interaction rather than on empirical research. The paper concludes with a discussion of an architectural, context-dependent theory of second language proficiency and its implications for the assessment of oral proficiency.

Introduction

The question that will be addressed in this paper is a simple one. It is, "What differences exist between those native-nonnative interactions in which the nonnative speaker (NNS) is a novice at interacting in the second language and those interactions in which the NNS is an experienced interactor?" The paper will examine two answers to that question that have been given in the language testing community. It will show that both of those answers are essentially unsatisfactory because they are based on mistaken intuitions about learners' interactional styles and an oversimplified understanding of second language proficiency. Another attempt will be made to answer the question by examining the research literature on native-nonnative interactions to see what

805

can be gleaned about proficiency-related differences in interactional style from empirical studies of NNS discourse. It will be argued that these studies show that oral proficiency in a second language is a multidimensional construct and that any description of oral proficiency needs to treat proficiency as an architecture of interrelated components rather than as a unitary construct. The architectural view of oral second language proficiency has important implications for theories of second language learning (cf. Young & Perkins, 1995) and yields valuable insights into ways of assessing oral proficiency.

Developmental Views of Oral Proficiency

On the surface at least, the question of what proficiency-related differences exist in native-nonnative interactional discourse is one that has already been answered to the satisfaction of many individuals, notably those who subscribe to the so-called "proficiency movement" in language testing and teaching. The *Proficiency Guidelines* published by The American Council on the Teaching of Foreign Languages (ACTFL, 1989), give extensive definitions of what is meant by speaking proficiency at nine different stages from "Novice-Low" through "Superior." Two of these stages are reproduced below. These guidelines have, however, been the subject of criticism by language testers and second language acquisition (SLA) researchers alike (Bachman, 1990; Bachman & Savignon, 1986; Barnwell, 1989; Lantolf & Frawley, 1985, 1988). These critics point out several defects in the *ACTFL Guidelines*. First, the *Guidelines* assume that proficiency in any skill is a unitary construct; second, they were not based on any empirical research on second language learners' linguistic performance; third, they have not been validated against the proficiency judgments of naive native speakers (NSs); and last, the *Guidelines* deal only cursorily with interactional discourse.

Two of the Nine Levels of the ACTFL Proficiency Guidelines *for Speaking*

Novice-High

> Able to satisfy partially the requirements of basic communicative exchanges by relying heavily on learned utterances but occasionally expanding these through simple recombinations of their elements. Can ask questions or make statements involving learned material. Shows signs of spontaneity although this falls short of real autonomy of expression. Speech continues to consist of learned utterances rather than of personalized, situationally adapted ones. Vocabulary centers on areas such as basic objects, places, and most common kinship terms. Pronunciation may still be strongly influenced by first language. Errors are frequent and, in spite of repetition, some Novice-High speakers will have difficulty being understood even by sympathetic interlocutors.

Intermediate-High

> Able to handle successfully most uncomplicated communicative tasks and social situations. Can initiate, sustain, and close a general conversation with a number of strategies appropriate to a range of circumstances and topics, but errors are evident. Limited vocabulary still necessitates hesitation and may bring about slightly unexpected circumlocution. There is emerging evidence of connected discourse, particularly for simple narration and/or description. The Intermediate-High speaker can generally be understood even by interlocutors not accustomed to dealing with speakers at this level, but repetition may still be required.
>
> (ACTFL, 1989, p. 1)

It is easy to show one of the problems of the unitary approach to oral proficiency taken by the authors of the *Guidelines*. The two ACTFL descriptions are of learners at the Novice-High and Intermediate-High levels and attempt to characterize learners' proficiency in several different areas. The first thing to notice is that the descriptions of each level contain several separate descriptions of different aspects of learner language at that level. So, at the Novice-High level, for example, learners' language is described in terms of syntax ("Can ask questions or make statements ..."), fluency ("Shows signs of spontaneity ..."), discourse features ("Speech continues to consist of learned utterances rather than of personalized, situationally adapted ones"), vocabulary ("Vocabulary centers on basic objects, ..."), pronunciation ("Pronunciation may still be strongly influenced by first language"), and comprehensibility ("some Novice-High speakers will have difficulty being understood even by sympathetic interlocutors"). On the basis of the descriptions in the ACTFL *Guidelines* therefore, it seems reasonable to conclude that as learners increase in proficiency, so each component of their interlanguage (IL) develops in step with all the others components. Thus, if we compare learners at the Intermediate-High and Novice-High levels, then, according to the *Guidelines*, the Intermediate-High learners not only know more vocabulary:

> *Novice-High:* Vocabulary centers on areas such as basic objects, places, and most common kinship terms.
>
> *Intermediate-High:* Limited vocabulary still necessitates hesitation and may bring about slightly unexpected circumlocution.

But they are also more comprehensible:

> *Novice-High*: Errors are frequent and, in spite of repetition, some Novice-High speakers will have difficulty being understood even by sympathetic interlocutors.

> *Intermediate-High:* The Intermediate-High speaker can generally be understood even by interlocutors not accustomed to dealing with speakers at this level, but repetition may still be required.
>
> (ACTFL, 1989, p. 1)

The problem with such an approach to second language spoken proficiency is that different components of interlanguage proficiency do not necessarily develop in step. For example, Beebe and Takahashi (1989) give a good illustration of how pragmatic competence may develop at a different rate from other interactional skills in English. Beebe and Takahashi note that the pragmatic competence of advanced and relatively assimilated Japanese ESL speakers may not be at the same level as their competence in other areas. Specifically, advanced Japanese learners of English are more direct in disagreeing and in giving embarrassing information than Americans feel is appropriate. This is perhaps because in an attempt to converge with American interlocutors, they go too far and end up by diverging from the patterns they perceive.

Such individual differences among learners cannot be handled by a proficiency system such as the ACTFL *Guidelines* because the underlying assumption is that all aspects of spoken second-language proficiency develop at the same rate. In the examples of vocabulary and pronunciation given above, how is one to assess a learner whose command of vocabulary fits the description at the Intermediate-High level but whose pronunciation fits the description at the Novice-High level? The problem is compounded when we consider that each description refers to at least six different components of a learner's proficiency and that, in principle, each component may be assessed at a different level from all the others.

Although assumptions of lockstep development of second language proficiency are inherent in the published ACTFL Guidelines, the practice of ACTFL raters has been to use the Guidelines as descriptions of threshold performance; that is, a certain rating is only given if all the components of language performance described at a given level are maintained. However, this practice is subject to criticisms similar to those made of the rating scale itself. First, what evidence exists that the particular combination of vocabulary, pronunciation, and pragmatic skills described by the Guidelines is a useful or even frequent characteristic of a particular stage in interlanguage development? In addition, the rating scale is set up as an implicational hierarchy, as is stated explicitly by Byrnes (1989), "Each major level subsumes the criteria of the levels below it" (p. 2-5). However, it appears from studies of task-related performance variation in oral proficiency interviews (e.g., Young, 1995; Young & Milanovic, 1992) that learners' interlanguage performance varies quite markedly with different tasks and with culture-specific understandings of the requirements of those tasks. Thus, criterion performance by a given learner on a particular OPI testing construct such as responding to a level check or probe does not necessarily imply either (a) that the same learner can perform similarly on other tasks, or (b) another learner who fails to reach criterion on this task will also fail on different tasks. Thus the ACTFL Guidelines as they are writ-

ten and as they are used, appear to be an inadequate characterization of proficiency differences among learners because they ignore the modularity and context-dependency of interlanguage proficiency.

One solution to the lack of modularity inherent in the ACTFL approach is to develop different scales for different components of proficiency. This is the approach taken by many modern examinations of oral skills, including the Cambridge Assessment of Spoken English (CASE) rating scales under development by the University of Cambridge Local Examinations Syndicate (UCLES). CASE raters use eleven separate scales, including seven entitled *Grammar, Vocabulary, Pronunciation, Organization, Communication Strategies and Interaction Skills, Interlocutor Support,* and *Task Achievement.* A letter score from A through E or from A through D is awarded on each scale. A rating on each scale is used to provide a profile of the learner's performance. Ratings on these scales are reported as a comment and not as a final score. The learner's final score is reported on the overall ability scale, which is arrived at independently of the sub-scales. As an example of this rating system, the CASE *Communication Strategies and Interaction Skills* scale is given below.

Five Levels of One of the Cambridge Assessment of Spoken English Rating Scales

Scale 5

(Communication Strategies and Interaction Skills)

A. Intended meaning communicated in all contexts of the assessment. Initiates and maintains interaction and responds to shifts of topic with ease.

B. Intended meaning communicated in most contexts. Initiates and maintains interaction and responds to shifts of topic with only occasional difficulty. Very few inappropriate utterances produced.

C. Main ideas communicated though repair strategies need to be used to overcome linguistic shortcomings. Experiences some difficulty when attempting to initiate or when responding to shifts of topic.

D. Communicates main ideas clearly in some contexts only. Repair strategies rarely used to overcome linguistic shortcomings. Rarely attempts to initiate interaction and experiences difficulty in responding to shifts of topic.

E. Experiences great difficulty in communicating and is unable to use repair strategies to overcome linguistic shortcomings. Many inappropriate utterances. Does not attempt to initiate and depends on others to maintain interaction.

(UCLES, 1991, p.24)

This approach to the assessment of oral proficiency is an improvement on the ACTFL system since it recognizes the modularity of interlanguage development, and deals much more explicitly with interactional discourse. However, the CASE system is subject to some of the same criticisms as the ACTFL scales. In particular, underlying the rating scale for each component of proficiency is the implication that language learning involves linear, monotonic increases in the component under consideration. As can be seen from the CASE *Communication Strategies and Interaction Skills* scale, for example, the descriptors are scaled on such a continuum.

Apart from the fact that "communication" is undefined in the CASE scale, evidently the view underlying the scale is that, whatever communication is, learners at higher points on the scale are better at it than those at lower points. The view that language learning involves linear monotonic increases in some variable—a one-directional vector if you like—is a common one and one that has some limited empirical support. For example, some vocabulary studies (Broeder, Extra, & van Hout, 1989; Broeder, Extra, van Hout, Strömqvist, & Voionmaa, 1988; Yorio, 1989) have described learners as gradually adding to their second language word-hoard, or store of conventionalized expressions. In addition, studies of learners' processing speed in second language comprehension tasks (Conrad, 1989; Griffiths, 1990) and one study of learners' speech rate (Lennon, 1990) have also shown a fairly steady increase in these variables over the time in which learners are exposed to the second language.

However, the vector mode of learning is the exception rather than the rule. A majority of longitudinal studies of interlanguage development have shown instead that the interlanguage system goes through a period of restructuring and reorganization. One result of such restructuring is that intermediate stages may be further from the target than either beginning or advanced stages—a pattern commonly called U-shaped behavior. A pertinent example of this nonlinear development in the field of interactional discourse is provided in an anecdote by Johanna Nichols (1988, p. 15).

> When I first traveled to the USSR, my classroom knowledge of Russian marked me as a foreigner, and interlocutors were uniformly supportive. When my fluency improved to the point that I was taken for a Baltic Soviet, expectations rose, support decreased, and communication became more difficult. This affected not only everyday communication but also scholarly interaction.

Nichols' anecdote illustrates that effective interactive communication involves not only the learner but also her NS interlocutors. NS interlocutors may be more tolerant of errors when the learner is at an intermediate stage than when she reaches an advanced stage of proficiency, making communication *less* effective at an advanced than at an intermediate level.

There are several SLA studies that have described the restructuring of interlanguage knowledge that occurs as proficiency develops. Most of the studies have to do with the acquisition of syntax. They include Huebner's (1985) description

of the development of a definite-article-like form *da* in the English interlanguage of his Hmong informant, and Meisel, Clahsen, and Pienemann's (1981) description of second language learners' cognitive reorganization in the acquisition of German word order. Ard and Gass (1987) also report that intermediate and advanced learners differ in the processes underlying their judgments of grammaticality. Intermediate learners appear to judge the grammaticality of sentences in a syntactically uniform manner, while advanced learners judge sentences differently according to the lexical items that they contain. While restructuring phenomena have received most attention from SLA researchers working in the field of syntactic development, there is also evidence of restructuring in learners' interactional discourse competence. This evidence is to be found in studies of discourse domains and communication strategies, which will be reviewed in the following section of this paper.

It therefore seems, from this brief review of developmental approaches to proficiency that our intuitive views of proficiency in interactive discourse can lead to some false conclusions regarding the differences between novice and expert patterns of interaction in a second language. First, as Bachman (1989) and Higgs and Clifford (1982) have previously suggested, proficiency does not necessarily develop at an equal rate in all its components and it is more reasonable to assume that development is modular rather than unitary. Second, proficiency does not necessarily develop in a linear fashion but may be subject to restructuring and reorganization. The rating scales that are part of oral proficiency tests such as the ACTFL oral proficiency interview and the Cambridge Assessment of Spoken English reflect such mistaken beliefs about learners' interactional behavior. As will become clear in the following section, when we look into actual examples of NS-NNS interactional discourse, learners' proficiency turns out to be a more complex construct than it appears intuitively.

The theoretical framework within which an alternative approach to oral proficiency will be developed is that put forward recently by researchers in the field of educational measurement (Snow, 1989, 1990; Mislevy, 1993; Young & Perkins, 1995). The framework is described in a recent paper outlining new developments in educational measurement, in which Mislevy (1993) describes the multidimensionality of learners' competence. According to Mislevy, competence is:

> a complex constellation of facts and concepts, and the networks that interconnect them; of automatized procedures and conscious heuristics; of perspectives and strategies, and the management capabilities by which the learner focuses his efforts. (p. 28)

One consequence of this multidimensionality is that:

> Increasing competence in a substantive area need not be reflected as uniformly increasing chances of success on all tasks. Patterns of smooth increase may be observed for certain people on certain sets of tasks, in certain phases of development. Discontinuous patterns

> of change begin to appear as the scope of tasks becomes broader, as the range of development becomes greater, and as the experience of [learners] becomes more diverse. (p. 30)

In the following section I will discuss evidence from studies of NS-NNS interactive discourse that show that second language learners' competence develops in the discontinuous and context-dependent way described by Mislevy.

Evidence for the Discontinuous Development of Oral Proficiency

Evidence that interactional oral proficiency is a multidimensional construct that develops in a discontinuous way can be found in two separate threads of SLA research. First, evidence will be reviewed of differential learner competence in different discourse domains; second, evidence will be reviewed suggesting that learners' strategic competence in interaction develops in a nonlinear fashion.

NNS Participation

The amount of participation in interactional discourse is one way in which we might expect expert NNS interactors to differ from novices. In the CASE rating scale, for example, there is a mention that learners at level *D* "rarely initiate," while learners at level *A* "initiate and maintain the interaction." Again, underlying the CASE scale is the view that advanced learners initiate and participate in conversations more than novices and this increasing participation is a result of increasing proficiency. However, such a neat correspondence between participation and proficiency is not born out by empirical research.

Several studies have shown that the nature of the content being discussed in NS-NNS interactions greatly affects learners' participation in the discourse and their grammatical accuracy. Studies by Zuengler (1989), Zuengler and Bent (1991), and Woken and Swales (1989) have shown that when both participants in a NS-NNS interaction believe that the NNS possesses greater expertise than the NS in the subject matter being discussed, then NNS participation increases and NNSs initiate more new topics than NSs. In a study of discourse in UCLES First Certificate in English oral interviews, Young and Milanovic (1992) found that the subject matter of the interview affected how long a given topic persisted throughout the interview. In addition, the emotional investment of NNSs in the subject matter under discussion has been shown to affect accuracy in interlanguage by Eisenstein and Starbuck (1989). As suggested by Zuengler (1992), these various findings appear coherent when they are interpreted in the light of Selinker and Douglas's discourse domain model of interlanguage performance (Selinker & Douglas, 1985; Douglas & Selinker, 1985). According to discourse domain theory, learners' competence may vary according to the content domain, or subject matter, of discussion. Thus, learners' grammatical accuracy and interactional style may depend on the domain. This clearly has important implications for judging

learners' proficiency. It is possible that learners at a generally low level of proficiency may appear to have considerable interactional competence when the topic of discussion is one that they appear to command. In contrast, learners at a generally high level of proficiency may have low interactional competence in domains that they do not command. And increases in interactional proficiency may occur as learners' knowledge of a domain increases, rather than because of a general increase in second language proficiency.

Communication Strategies

A second way in which we might expect expert and novice second language interactors to differ is in their use of communication strategies. Communication strategies are defined by Tarone (1981, p. 288) as those attempts by speakers and their interlocutors "to agree on a meaning in situations where requisite meaning structures do not seem to be shared." From the CASE rating scale, it appears that use of repair strategies increases from level *E* through level *C*. However, studies have shown that such a picture is oversimplified and that in fact the learner's use of communication strategies undergoes considerable restructuring and reorganization as proficiency increases.

Poulisse (1990) noted that there was indeed an inverse relationship between the number of compensatory strategies used and the proficiency level of Dutch subjects learning English as a foreign language. She suggested a straightforward explanation of this fact, namely that low proficiency learners "encounter more lexical problems and therefore need to resort to compensatory strategies more often" (p. 145). However, she also noted that, on some tasks, high proficiency learners use different kinds of communication strategies from low proficiency learners. High proficiency learners tend to use more holistic conceptual strategies, typically involving semantic approximation by means of words or gestures. On the other hand, low proficiency learners tend to use more transfer-based strategies. Again, the explanation for these proficiency-related differences is quite straightforward—"learners of a lower proficiency level do not have a sufficiently large L2 vocabulary at their disposal to come up with suitable approximations" (Poulisse, 1990, p. 145).

In another detailed examination of proficiency-related differences in the communication strategies of Persian learners of English, Paribakht (1985) found that both the type of strategy and the overall frequency of strategy use vary with learners' proficiency in ways similar to those established by Poulisse. However, both Paribakht (1985) and Bialystok (1990) note that strategic competence is not something that develops in a second language but is instead an ability that develops first in a child's acquisition of its first language. Both NSs and NNSs use of communication strategies and proficiency-related differences in NNS strategy use are not due to differing underlying competence, but rather due to (a) different kinds of communication difficulty and (b) different linguistic resources that individuals can use in order to overcome those difficulties.

The proficiency-related differences in kind and number of strategies noted above were also observed in a study of strategic behavior of Japanese learners of English based on Foreign Service Institute oral interviews by Yoshida (1991). Yoshida's study compared strategic use by learners at three proficiency levels: beginners, intermediate, and advanced. She found that intermediate-level learners' strategy use differed significantly from *either* high *or* low proficiency learners. She found that intermediate learners use more reduction strategies than either beginners or advanced learners. Reduction strategies involve changing or abandoning communication goals. The reason for the greater use of this kind of strategy at the intermediate level, Yoshida suggests, is that intermediate learners begin to use more complex syntactic forms and this leads NS interlocutors to broach topics that the learners cannot handle. Since the learners cannot communicate effectively on such topics they are forced to abandon them. However, NS interlocutors do not attempt difficult topics with beginners since the NSs know that the learners cannot handle them. On the other hand, high proficiency learners' control of greater linguistic resources enables them to deal with difficult topics by means of achievement strategies.

This review of proficiency-related differences in communication strategies suggests that strategic competence involves allocating whatever communicative resources a speaker commands in order to overcome communication difficulties. Speakers obviously differ in the nature and amount of communicative resources that they can call on—and as a result speakers at different levels of proficiency use different strategies. However, second language learners do not simply use more strategies at higher levels. Strategy use does not follow the linear, monotonically increasing vector that underlies conventional rating scales. This ability to deploy communicative resources to overcome communication difficulties originally develops, as Bialystok (1990) suggests, in children as they learn their first language. Later, when children or adults come to learn a second language, they already have experience of overcoming communication difficulties and have developed certain strategies for doing so. In other words, beginning learners of a second language already have strategic competence. As learners' knowledge of a second language develops, however, the kinds of challenges that learners at different proficiency levels face in interacting with NSs are very different. In particular, as we have seen, the choice of topic by NSs is influenced by their perceptions of their NNS interlocutors' proficiency level. Because learners at different proficiency levels control different syntactic and lexical resources, their strategies for overcoming communication difficulties are also different at different levels. Learners do not become more strategically competent at higher proficiency levels. Rather, as proficiency develops, learners use different strategies to allocate whatever communicative resources they control to overcome communication difficulties.

Conclusions

This paper began with the question, "What differences exist between those native-nonnative interactions in which the NNS is a novice at interacting in the second

language and those interactions in which the NNS is an experienced interactor?" I hope to have shown that the answers to that question that are based on intuition rather than research are mistaken. In particular, the ways in which differences in oral proficiency are described in two recent testing instruments are unsatisfactory because (a) the ACTFL *Guidelines* ignore the modularity of linguistic competence, and (b) the ACTFL and CASE rating scales assume a linear, monotonic increase in competence in all components of proficiency.

When we turn from speculation to examine the empirical studies of differences in oral proficiency, we see that the latter depends on discourse domain as much as it depends on lexical and syntactic knowledge and that some components of oral proficiency such as communication strategies develop in a highly nonlinear way, as learners learn to adapt their limited linguistic resources to overcoming the difficulties they face in communication.

These descriptions of how learners become experts at interacting with NSs reveal that experts do not simply have more of what novices lack but that the factors underlying their proficiency are different and interact in different ways. As Bialystok (1991, p. 75) has put it,

> Language proficiency is not a single achievement marking some quantitative level of progress with language learning. Rather, it is the ability to apply specific processing skills to problems bearing identifiable cognitive demands. Proficiency in a domain, or in a task, is evident when the demands of the task are not in excess of the demands of the language learner. Thus language learners with a particular configuration of skill component development will in fact exhibit a range of proficiency with the language that is determined by the impact of the task demands on the processing abilities of the learner.

This view of second language proficiency as a modular, highly context-dependent construct, whose components develop in possibly discontinuous spurts and stages poses serious problems for assessment. One radical possibility may be that assessment instruments must be so varied as to thwart all attempts to come up with bases for comparing language performance by speakers in different contexts. Another, more optimistic outcome may be that we start from the premise that learners' competence involves a complex constellation of knowledge, skills, and the networks that interconnect them. An assessment of that competence at a given moment in time thus means some kind of map of that constellation. If we then are able to carry out further empirical spadework which allows us to characterize the different trajectories by which competence changes, we may be in a better position to predict future performance in a wide range of contexts on the basis of present performance in a single context.

References

American Council on the Teaching of Foreign Languages. (1989). *ACTFL proficiency guidelines*. Yonkers, NY: Author.

Ard, J., & Gass, S. M. (1987). Lexical constraints on syntactic acquisition. *Studies in Second Language Acquisition, 9*, 233-252.

Bachman, L. F. (1989). Assessment and evaluation. *Annual Review of Applied Linguistics, 10*, 210-226.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. New York: Oxford University Press.

Bachman, L. F., & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal, 70*, 380-390.

Barnwell, D. (1989) 'Naive' native speakers and judgements of oral proficiency in Spanish. *Language Testing*, 6, 152-163.

Beebe, L. M., & Takahashi, T. (1989). Do you have a bag? Social status and patterned variation in second language acquisition. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition* (Vol. 1: *Discourse and pragmatics*), 103-125. Clevedon, Avon: Multilingual Matters.

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Oxford: Basil Blackwell.

Bialystok, E. (1991). Achieving proficiency in a second language: A processing description. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogy: A commemorative volume for Claus Faerch*, 63-78. Clevedon, Avon: Multilingual Matters.

Broeder, P., Extra, G., & van Hout, R. (1989). Processes in the developing lexicon of adult immigrant learners. *AILA Review-Revue de l'AILA, 6*, 86-109.

Broeder, P., Extra, G., van Hout, R., Strömqvist, S., & Voionmaa, K. (1988). *Processes in the developing lexicon*. Strasbourg: European Science Foundation.

Conrad, L. (1989). The effect of time-compressed speech on native and EFL listening comprehension. *Studies in Second Language Acquisition, 11*, 1-16.

Douglas, D., & L. Selinker. (1985). Principles for language tests within the 'discourse domains' theory of interlanguage: Research, test construction and interpretation. *Language Testing, 2*, 205-226.

Eisenstein, M. R., & Starbuck, R. J. (1989). The effect of emotional investment on L2 production. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition* (Vol. 2: *Psycholinguistic issues*), 125-137. Clevedon, Avon: Multilingual Matters.

Griffiths, R. (1990). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning, 40*, 311-336.

Higgs, T. V., & Clifford, R. T. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher*, 7-79. Skokie, IL: National Textbook.

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning, 35*, 141-164.

Lantolf, J. P., & Frawley, W. (1985). Oral proficiency testing: A critical analysis. *Modern Language Journal, 69*, 337-345.

Lantolf, J. P., & Frawley, W. (1988). Proficiency: Understanding the construct. *Studies in Second Language Acquisition, 10*, 181-195.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387-417.

Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*, 109-135.

Mislevy, R. J. (1993). Foundations of a new test theory. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests,* 19-39. Hillsdale, NJ: Erlbaum.

Nichols, J. (1988). Language study, international study, and education. *Profession, 88*, 10-17.

Paribakht, T. (1985). Strategic competence and language proficiency. *Applied Linguistics, 6*, 132-146.

Poulisse, N. (1990). *The use of acquisition strategies by Dutch learners of English.* Dordrecht, Holland: FORIS.

Selinker, L., & Douglas, D. (1985). Wrestling with 'context' in interlanguage theory. *Applied Linguistics, 6*, 190-204.

Snow, R. E. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher, 18*, 8-14.

Snow, R. E. (1990). New approaches to cognitive and conative assessment in education. *International Journal of Educational Research, 14*, 455-473.

Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly, 15*, 285-295.

University of Cambridge Local Examinations Syndicate. (1991). *Cambridge Assessment of Spoken English: Assessor document.* Cambridge: Author.

Woken, M. D., & Swales, J. (1989). Expertise and authority in native-non-native conversations: The need for a variable account. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition* (Vol. 1: *Discourse and pragmatics*), 211-227. Clevedon, Avon: Multilingual Matters.

Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. K. Obler, (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss,* 55-72. New York: Cambridge University Press.

Yoshida, Y. (1991). *The use of communication strategies by adult Japanese speakers of English as a second language in language proficiency interviews.* Unpublished master's thesis, Southern Illinois University at Carbondale.

Young, R. (1995). Conversational styles in language proficiency interviews. *Language Learning, 45*, 3-42.

817

Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition, 14*, 403-424.

Young, R., & Perkins, K. (1995). Cognition and conation in second language acquisition theory. *IRAL, 33*, 176-197.

Zuengler, J. (1989). Performance variation in NS-NNS interactions: Ethnolinguistic difference, or discourse domain? In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in second language acquisition* (Vol. 1: *Discourse and pragmatics*), 228-244. Clevedon, Avon: Multilingual Matters.

Zuengler, J. (1992). *Explaining NS-NNS interactions: Too many theories, or too few?* Paper presented at the annual conference of the American Association for Applied Linguistics, Seattle, WA, February 28-March 2.

Zuengler, J., & Bent, B. (1991). Relative knowledge of content domain: An influence on native-non-native conversations. *Applied Linguistics, 12*, 397-415.

Author

RICHARD YOUNG, Associate Professor and Director, Program in ESL, Department of English, University of Wisconsin-Madison, 600 North Park Street, Madison, WI 53706. Specializations: face-to-face interaction, interlanguage variation, language proficiency interviews.

Applied Language Learning
1995, Vol. 6, Nos. 1 & 2, pp. 27-39

Silence, Interruptions, and Discourse Domains: The Opportunities to Speak

Raimundo Mora
Rutgers University

Production-based theories that emphasize the role of speaking in second language acquisition (Swain, 1985; Ellis, 1990, 1986, 1987; Long, 1983) have encouraged language teachers to give students more opportunities to speak. Earlier, in a study on the pragmatic aspects of language use in ESL classes (1992), I researched possibilities of creating these opportunities. In this article, I discuss characteristics of a teacher language that give students opportunities to articulate ideas in an academic discourse as well as factors that prevent them from using such opportunities. I use the data to illustrate how some types of teacher silence can create opportunities for students to participate in classroom discussions, and how matching the content of the discussions with student culture and familiarity with the topic determine both participation and interruptions.

Method

In my analysis, I applied speech act theory (Austin, 1962), a branch of pragmatics, because it clarifies the differences between the grammatical and pragmatic competencies that speakers need to express their intentions effectively and appropriately (Leech, 1983). I integrated speech act theory with discourse analysis in the tradition of ethnography of communication (Hymes, 1972; Gumperz, 1982) that allowed me to document participant interpretations of the language of the discussions. I documented language use in three classrooms by observing, taking notes, and videotaping class sessions that included discussions and lessons. Discussions were organizations of talk in which students explored topics that allowed them to develop a more sustained discourse. Lessons, on the other hand, followed an interactional format in which teachers knew the answers and systematically evaluated them as right or wrong. Consequently, lessons did not give students the same opportunities for discourse. Therefore I chose discussions for microanalysis. Afterwards I watched the videotapes with the participants and tape-recorded their interpretations.

In accordance with the ethnographic method used by Saville-Troike (1982), when participants commented on aspects of language use, I asked teachers and stu-

©1995, Raimundo Mora

819

dents for more specific information and then compared their answers to the same questions. Subsequently, I transcribed class discussions and repeatedly compared different versions of the same discussions: videotapes, transcripts, participant interpretations, and the logs. I organized the transcripts into speech acts and numbered them consecutively for each class discussion, and indicated the time elapsing between them. In the transcripts I also included overlaps and nonverbal clues as they were interpreted by participants and the researcher.

Subjects

I conducted the study at a community college in a large metropolitan area. This college, like other community colleges, seeks to meet the needs of its community. Nine thousand students were enrolled in the credit programs. At the time of the study, 40% of the student population consisted of Limited English Proficiency (LEP) students, and one of the main functions of the college was to provide them with educational programs. The ethnic composition of the student body was as follows: 18% Caucasian; 30% Hispanic; 29% Afro-American; 13% Asian; and 6% other. These percentages reflect the demographics of the college community. Sixty percent of these students live in the neighborhood of the college, which is one of the largest neighborhoods in the city and one of the most ethnically diverse.

The subjects of my study were 3 ESL native English teachers and 23 Spanish-speaking students. To protect their privacy, I refer to the teachers as Justine, Janis, and Joe. I have also renamed the students. Students placed in the three ESL classes that I studied were at an intermediate level. There were five levels in the ESL program that offered these classes, two levels below the intermediate level and two above it. In each of the three classes, I spent one term collecting data.

Discussion

Teacher Silence

Silence as a means of giving students opportunities to speak has not been researched except for a few studies about "waittime," the time elapsing between the teacher questions and the student answers, and between the student answers and the following teacher speech act. Rowe (1986) found that it is common for teachers to wait 1 second or less after asking a question for a student to start a reply, and less than a second to react or ask the next question after a student stops speaking. She also found that when teachers waited more than 2 seconds, the students were more likely to elaborate on their answers. Unlike Rowe, who examined the silent interlude between questions and answers, I looked at every period of silence lasting 2 or more seconds during selected classroom discussions. I observed that each of the teachers had different ways of using silence to create opportunities for student participation. I identified nine kinds of silence and found that some created spaces for students to talk and some did not.

Types of Silence that Gave Students Opportunities to Initiate Interactions

1. *Silences after Student Statements* (The teachers waited more than two seconds for students to initiate an interaction.)

Extract 1.1

Janis (Teacher)	Did you learn to cook by yourself or did somebody teach you?
Francisco (Student)	Alone. (*2 seconds*)
Rosa (Student)	Alone? That's good.
Francisco (Student)	I saw my mother cooking, and I learned.
Dora (Student)	The mothers have to teach. The mothers have to teach their sons to cook.
Rosa (Student)	Yes. (*2 seconds*)
Dora (Student)	And...and care everything....(*2 seconds*) It is very important for their future.

2. *Silences after Teacher Statements, Comments, and Commands* (The teachers waited more than two seconds for students to initiate an interaction.)

Extract 1.2

Joe (Teacher)	Money is power. It's true between countries. It's true between cities; it's true in relations. You have the money. You have the power. And 10% of the people own half the money. (*10 seconds*)
Maria (Student)	They have the power. Do they pay taxes? (*Laughter in the back.*) They chose the president and make and make and break the law.

Types of Silence that Gave Students Time to Answer Questions and Finish an Utterance

3. *Silences after Student Questions* (The teachers waited more than two seconds after student questions.)

Extract 2.1

Rosa (Student)	He has to learn to cook to get married, right? (*laughs*) (*6 seconds*)
Monty (Student)	A lot of people don't have support; they don't have a house; they don't have nothing, but they have to marry.

4. *Silences after Teacher Questions* (The teachers gave students more than two seconds to answer their questions.)

Extract 2.2

Joe (Teacher)	Why do you think there are so many homeless children in New York? (*4 seconds*) Have you ever thought of that? Have you ever wondered why? (*24 seconds*)
Musako (Student)	Are these children citizens?

5. *Teacher Silence during Student Pauses* (The teachers did not interrupt students when students paused for more than two seconds in the middle of an utterance.)

Extract 2.3

Joe (Teacher)	I would say that most of them are citizens.
Maria (Student)	Some of these children come from. . . (*4 seconds*) immigrants.

6. *Scaffold Silences* (When the students did not know the answer to a question, the teacher provided it, but stopped halfway so that the student could finish the utterance.)

Extract 2.4

Joe (Teacher)	If 10% of the people own 50% of the money, I think we all agree that in this country money equals. . . (*3 seconds*)
Maria (Student)	Because [in]
Joe (Teacher)	wait, wait. The word I was looking for was "power."

Types of Silence that Did not Provide Students with an Opportunity to Speak

7. *Teacher Pauses* (Teachers paused for more than 2 seconds in the middle of an utterance to emphasize what they had just said or to gather their thoughts. These silences were not intended to be ended by the students.)

Extract 3.1

Janis (Teacher)	So whoever can afford it pays for the wedding pays for it and that's OK. (*8 seconds*) And. . . . Is it true what he said that all women want to be married?

8. *Ritual Silences* (The teacher remained silent for more than two seconds after the first utterance that he or she used to start a discussion in order to get the student attention.)

Extract 3.2

Janis (Teacher)	OK. (*5 seconds*) What would you tell a 20-year- old who comes to you for advice? Should they consider marriage ever or shouldn't they consider marriage?

9. *Action Silences* (when the teacher who held the floor stopped talking in order to perform a nonverbal action, e.g. erasing the blackboard).

Extract 3.3

Justine (Teacher)	Let's talk about the third eye (*4 seconds*) (*writing "third eye" on the blackboard*)

The Impact of Teacher Silence on Student Participation

In Joe's discussion, eight of his silences opened spaces for student talk. There were two instances in which he waited for the students to answer his questions, four in which he waited for students to finish utterances, and two scaffold silences. Only three of his silences gave students the opportunities to initiate an interaction: two silences after student statements and one after one of his statements.

In Justine's discussion, three of her silences opened spaces for student talk. There were two silences that came after her questions, which were not ended by the students, and one silence when a student stopped in the middle of an utterance to gather her ideas. None of her silences gave students an opportunity to initiate an interaction.

In Janis's discussion, 24 of her silences opened spaces for student talk. There were 7 instances in which she gave students time to answer her questions. Also there were 2 instances in which she gave students time to finish their utterances, and 1 in which she gave students time to answer each other questions. Finally, there were 14 instances in which she gave students opportunities to initiate interactions, 12 after student statements and two after her statements.

Thus Janis gave students more opportunities to speak than the other two teachers combined and her students participated in the discussion more than the students in the other two discussions combined. In her discussion, most instances of her silence after student statements (Silence No. 1) were ended by students who initiated interactions. This type of silence promoted student-to-student interactions because when students initiated interactions, they frequently addressed each other. Every silence after her statements and commands (Silence No. 2) was also ended by the students. The students spontaneously initiated interactions related to their different social and cultural backgrounds.

Janis used silence in class consistently with her reported intentions. She explained that she tried to give them both time and information because LEP students needed more time than native speakers to organize their ideas and information about the topics discussed in class. Indeed, she not only conducted discussions at a slow pace and frequently used kinds of silence that helped students prepare answers and initiate interactions, but also made sure that students received information through handouts and movies prior to class discussions. I discuss the impact of this strategy on student participation in the *Discourse Domains* section.

Discourse Domains

Although teacher silence after posing questions was supposed to give students enough time to answer, frequently students did not answer these questions. It seems that if students did not know the answers, extra time did not help them. Likewise, students who were not familiar with the themes of the discussion could not ini-

tiate interactions even if they had opportunities. Also, students whose cultural norms were in conflict with the topics remained silent. However, it seems that when students initiated interactions, they had more opportunities to find something to say than when they did not know the answer because they could bring their familiar domains into the discussions. In approximately half of the cases, students did not answer the teacher questions, and the teacher performed the next utterance either by rephrasing questions or issuing a command. However, silence after teacher statements and commands and after student statements was ended by the students in 72% of the cases.

According to Selinker and Douglas (1985), individuals who are learning a language create discourse domains of various slices of life in which they can be articulate. The authors documented how the same speakers could be articulate in a familiar context and inarticulate in an unfamiliar context. In the analyzed discussions, ability of students to participate was determined in part by their familiarity with the subject. During a discussion that took place after Justine showed the film *Psycho*, Norberto and Marina spoke at length about the double personality of Norman Bates, the main character in the movie. When Justine tried to end the discussion, these two students continued talking to her about the movie. She had to make two more attempts and promise to continue the discussion the next day. They impressed the teacher with their comments because they had background information about the psychological issues the film addressed. Namely, Norberto had taken an introductory psychology course when he attended law school in Argentina, as had Marina when she attended college in Ecuador. Thus they were aware of the psychological concepts they needed to express themselves. In the class discussion, they asked and answered questions, recounted episodes from the movie, and provided explanations. By contrast, the other students did not speak at all or limited themselves to one or two short statements that revealed the fact that they had not understood the movie. One of these students was Bertha, who attempted to explain the movie by saying that it was about a madman and his mother, when in reality the son and the mother were the same person. When I asked her to summarize the plot of the movie in Spanish, she said that it was about a mother and a son who ran a hotel and that the mother had killed a female guest out of jealousy. Although Bertha had understood one level of the plot, she missed the last important twist of the film. Despite the fact that Bertha always looked for opportunities to speak, she frequently ran into sociocultural obstacles that prevented her from engaging the teacher in sustained interactions. Almost all of her speech acts were clarification questions. Thus Bertha's ability to talk about *Psycho* was hindered by her lack of information about the psychological issues and her unfamiliarity with film narrative.

I also found that in addition to familiarity with the subject, cultural and circumstantial factors determined the domains in which students could be articulate. For example, Janis had students read an article about AIDS. After explaining new vocabulary, she opened a discussion about the AIDS crisis in the United States. At the end of the discussion, she asked students to write an essay about this topic for the next class. Rosa, who was from a small Ecuadorian village, was usually a lively participant, but in this class she looked distressed and did not speak. During the analysis of the videotape,

she said that she did not like to talk about AIDS, and added that "women should not talk about sex." I could not get her to say anything more about it. Rosa never turned in the composition on AIDS. Her difficulties speaking about the topic may have been personal, but her remark that women should not talk about sex also acknowledges gender differences in her native culture that played a role in her class performance. Marvin, who usually participated very actively in class, also remained silent during the discussion. In the analysis of the videotape, he said that his close friend had recently died of AIDS, so it was painful for him to talk about it in class. The emergence of the theme of pain took me by surprise. After their interpretations, I wrote in my log that in addition to the tension between linguistic proficiency and background knowledge about AIDS, student participation in class was also influenced by cultural constraints and personal circumstances.

In addition to unfamiliarity with the themes of the discussions, unfamiliarity of the students with the proper register created difficulties. In all three classes, the students frequently resorted to their experiences, which correspond to what Selinker and Douglas call the "life-story" domain, in response to teacher questions. Such questions were frequently phrased in an academic discourse, corresponding to what Selinker and Douglas call "specific-purpose" domain. This disparity frequently led to communication breakdowns. For example, before starting a discussion of poverty, Joe gave students some statistics about poverty in New York City and then asked for its causes. Maria related her own experiences of poverty in the city without making explicit connections to the causes of poverty. Joe, who did not understand her point, asked a clarification question, but could not understand her answer so he gave up and posed a question to another student.

In addition to Maria's use of a register that was not expected by the teacher, lack of contextualization of her life-story performances contributed to communication breakdowns with the teachers. Joe had to watch the videotape of the discussion several times to understand the relevance of Maria's story to his question about the causes of poverty. Subsequently, he realized that he had not understood some of her answers in part because she had not given him a "straight answer" and because she had not mentioned the connection between her stories and the questions he had posed. Maria's difficulties in expressing her intentions in class contrasted with her fluency in speaking Spanish during the videotape analyses. When she spoke Spanish with me, she often articulated her points of view by telling stories that she connected clearly to the topics of discussion. Apparently, she tried to do the same when she spoke English, except that she did not relate her stories overtly to the topics. Kramsch (1991) claims that LEP students need to learn not only how to use language in different contexts, but also how to use language as context. She explains that to communicate intentions effectively, LEP speakers have to assess the information that the listeners need. Speakers need to use language to create contexts within which their utterances are understood.

The discrepancy between the register that students used and the register expected by the teacher created communication conflicts that the teacher solved to promote student participation. When Janis asked her students to play the role of marriage

counselors and tell the class how they would advise somebody who was about to get married, she did not succeed in having students perform the role of marriage counselors, even though they were all familiar with the concept of marriage. They had been exposed to marriage in their own families and communities. Moreover, some of them were married. Nevertheless, they found it difficult to play the role of a marriage counselor. Because it was an unfamiliar role for some students, they defined the situation in terms of a familiar context: a conversation among peers about marriage. Under these circumstances, Janis did not object to this change of direction in the discussion. She allowed students to talk to each other within the realm of familiar, which led to their performance of a large variety of speech acts. Had she insisted that they play the role of counselors, she would not have elicited the same degree of participation, unless she had also provided instruction on the work of marriage counselors and the language needed to advise clients.

Interruptions

In this section, I analyze instances in which teacher interruptions limited student opportunities to speak, as well as instances in which such interruptions created communication conflicts and prevented students from articulating their ideas. To categorize interruptions and overlaps in the discussions, I used Coates's (1990) definitions of these terms. He uses the term "interruptions" for acts of seizing someone's turn and "overlaps" to describe instances of slight over-anticipation by the next speaker.

In a class discussion, Justine asked everybody questions about the concept of the "third eye." The discussion ended when all the students had answered. On five occasions, to ensure that all the students would have an opportunity to speak, she interrupted those who wanted to elaborate on their answers. Thus Justine interrupted her students frequently to allocate the next turn.

Joe interrupted the students to correct and clarify their performances. He did not tolerate student interruptions or overlaps. On two occasions, he stopped talking because Maria spoke at the same time as he did. When Joe analyzed this event, he said, "She used to interrupt me. We had an uneven communication." However, he did not comment on his own interruptions. In a separate session, Maria claimed no awareness of interruptions by either side.

Maria and Joe thought that it was appropriate for a teacher to interrupt a student. However, Joe found it inappropriate for a student to interrupt him. His interpretation of "interruptions" implied discreet functions for teacher and student registers. If it was a common principle of classroom behavior in this college that students did not speak at the same time as the teacher, then Maria's unawareness of or disregard for this principle probably contributed to her difficulties in communicating with the teacher.

It is noteworthy that when Joe said that Maria interrupted him all the time, he was referring not only to her overlaps but also to her obstructions of some of his attempts to reestablish the teacher-question/student-answer pattern of his discussions. For example, there was an instance when Joe asked a question and Maria tried to add

827

a comment to her answer; he interrupted her with yet one more question. This exchange initiated a sequence of five questions, all of which Maria answered. Joe ended the interaction with a statement that summarized his understanding of what Maria had said. Thus his questions effectively prohibited her from developing a topic beyond answers to the teacher's questions.

The disparity in the way teachers and students used and interpreted simultaneous talk can be explained in part by their cultural differences. Maria's use of simultaneous talk is consistent with practices of many Spanish-speaking communities, in which such behavior is considered appropriate and enthusiastic. This finding also corroborates Tannen's claim (1984) that different interpretations of overlaps and interruptions by different speech communities are frequent causes of communication conflicts and breakdowns.

In Janis's class, on the other hand, neither side spoke at the same time as the other during the teacher-student interactions. Janis tended to remain silent after the students finished talking rather than to interrupt them. There was only one instance in which Janis overlapped to take the floor away from the students. Otherwise her overlaps were meant to encourage students to keep the floor. However, the students did speak simultaneously in their interactions with each other. This pattern indicates the existence of at least two distinct registers for the students: one to use with the teacher and one to use with each other. Janis's students overlapped with each other in two different ways: to take the floor and to show cooperation.

In college discussions documented by Kuhn (1984), which took place in content classes and involved native English speakers, overlaps and interruptions were frequent. Kuhn attributes this frequency to the fact that students had to compete for the floor. He found that the number of interruptions by students was proportionate to the class size. He explains that in larger classes, the students had to compete harder with each other for the floor. They had to use interactive skills that included overlapping with other speakers at the right moment to get the opportunity to speak.

Students taught by Joe and Justine used language in ways that Kuhn (1984) and Cazden (1988) associate with lessons rather than discussions. Only students in Janis's class used interruptions and overlaps in a way comparable to the students in Kuhn's study. When Joe's and Justine's students interrupted teachers, they usually created communication conflicts that prevented them from having sustained interactions. Van Lier (1988) stated that interruptions for the sake of repair or assistance do not constitute floor threats in the same way that they might in general conversation. I would specify that these interruptions might not be a threat to the students when performed by the teacher, but they might indeed be a threat to the teacher if performed by students.

Conclusion

In this article I have focused on a few of the many pragmatic aspects of class discussions, namely silence, discourse domains, and interruptions. The data show the complex factors involved in cross-cultural communication in ESL classes and point to

ways in which teachers can promote class participation with implications for second language acquisition. The production-based theories cited at the beginning of this article claim that the more oral interactions learners have, the more they can test their linguistic hypotheses, and, consequently, the more language they acquire. According to the topicalization hypothesis, when the learner has some control over the topic, there is more classroom acquisition (Ellis, 1990). Ellis explains that when learners are given opportunities to propose topics of discussions, they are allowed to use their existing resources in communication.

Out of nine types of silence that I identified in this study, four gave students time to answer questions and finish utterances and two gave students opportunities to initiate interactions and therefore introduce their own topics. Teacher silence was often the key to student participation. A variety of factors, however, seemed to hinder this participation. Some discussion topics were too unfamiliar, painful, or inappropriate for students to talk about; sometimes teachers expected their students to use unfamiliar registers. The rules for interruptions seemed to vary from student to student, student to teacher, and teacher to teacher. Some of the variations in interruptions had an impact on student participation in the discussions. Interruptions by teachers could prevent students from developing a sustained discourse. Interruptions by students had different effects on participation, depending on whether they interrupted the teacher or each other.

Implications

These findings point to ways teachers can create spaces for students to participate in class discussions. For example, they can promote language use and acquisition by selecting appropriate topics, using types of silence that promote student participation, and avoiding interruptions. Closer attention to the content of classroom discussions and student discourse domains could gradually bring the two together. If students can perform only in life-story domains, teachers should encourage them to draw on these domains, while providing the linguistic and sociocultural instruction to articulate their ideas in a more abstract academic discourse. Teachers can also introduce students to various registers, focus on strategies that help students to express their ideas clearly, and explain cultural differences that might cause misunderstandings.

It was clear for all those involved in the study that to teach students English in an academic discourse was the main purpose of classroom discussions. However, the teachers were frequently caught in a contradiction between their reported pedagogical intentions—mainly, to have the students engage in an academic discourse—and what they actually did in class. During the analyses of the videotapes, it became apparent that they frequently did not bring their language use and their ideology into accord. Moreover, the teachers gave the impression that to a large extent they improvised when they elicited student talk, in contrast to their careful understanding and preparation of reading, writing, and grammar lessons. I have found that sometimes the use of language in the classroom did not match teacher intentions because they were not aware

of some aspects of language use. This finding is true also for language use in other settings. Speakers use language without being aware of its pragmatic aspects. Hall (1959) claims that language is culture and that culture controls behavior in ways removed from awareness. If this is true, then teachers have to make an effort to become aware of the ways in which they use language in the classroom. The use of videotapes to analyze teacher classroom performance might be one way to promote such awareness.

I found that the analysis of videotaped classes was a rewarding experience for everyone involved. The students and teachers had an opportunity to double-check their original interpretations of classroom situations and to discover new layers of meaning. I learned to learn from their interpretations. Thus videotapes can be a useful improvement tool for ESL teachers and students. An outside observer might not always be necessary. Teachers could videotape their classes, and then sit down with their students to sort through what went on. During the last decade the notional-functional approach has developed which emphasizes teaching the communicative aspects of English. However, the authors of textbooks and materials based on this approach emphasize linguistic formulas that native speakers use to perform different linguistic functions. Unfortunately, they do not tell us much about how teachers and students negotiate meaning in class.

Studies that document how language is used in ESL classes can help teachers to achieve their academic goals. Silence, interruptions, and discourse domains are only a few of many pragmatic aspects of classroom communication. Other issues, such as questions, effective language, humor, irony, and discourse strategies deserve more attention.

References

Cazden, C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann Educational Books.

Coates, J. (1986). *Women, men, and language*. New York: Longman.

Ellis, R. (1986). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (Ed.) (1987). *Second language acquisition in context*. Englewood, NJ: Prentice Hall International.

Ellis, R. (1990). *Instructed second language acquisition*. Cambridge, MA: Basil Blackwell.

Gilmore, P. (1985). Silence and sulking: Emotional displays in the classroom. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence*. Norwood, NJ: Ablex.

Gumperz, J. (1982). *Discourse strategies*. New York: Cambridge University Press.

Hall, E. (1959). *The silent language*. New York: Doubleday.

Hymes, D. (1972). Introduction: Toward ethnographies of communication: The analysis of communicative events. In P. P. Giglioli (Ed.), *Language and social context*. Middlesex, England: Penguin.

Khun, M. (1984). *A discourse analysis of discussions in the college classroom.* Unpublished doctoral dissertation, Harvard University, Cambridge.

Kramsch, C. J. (1991). *The challenge of the learner: Learning language in context or language as context? Plenary Address at the Annual Meeting of the American Association of Applied Linguistics*, New York City.

Leech, G. (1983). *Principles of pragmatics.* New York: Longman.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of meaning. *Applied Linguistics, 4.*

Mora, R. (1992). Pragmatic aspects of the language used in three selected ESL classes. Unpublished dissertation. New York University.

Rowe, M. (1986). Wait time: Slowing down may be a way of speeding up. *Journal of Education, 37.*

Saville-Troike, M. (1982). *The ethnography of communication: An introduction.* New York: Basil Blackwell.

Selinker, L. & Douglas, D. (1985). Wrestling with "context" in interlanguage theory. *Applied Linguistics, 6* (2).

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In M. Gass, & C. Madden (Eds.), *Input in second language acquisition.* Rowley, MA: Newbury House.

Tannen, D. (1984). *Conversational style: Analyzing talk among friends.* Norwood, NJ: Ablex.

Van Lier, L. (1988). *The classroom and the language learner.* New York: Longman.

Author

RAIMUNDO MORA, Assistant Professor, Academic Foundations Department, Rutgers University, 85 Columbia St, Apt. 21C, New York, NY 10002.
Specializations: sociolinguistics, pragmatics, English as a second language.

Applied Language Learning
1995, Vol. 6, Nos. 1&2, pp. 41-64

Pragmatics in Interlanguage: German Modal Particles

Martina Möllering
Macquarie University

David Nunan
University of Hong Kong

The comprehension and correct use of German modal particles, which occur most frequently in dialogues, are indicators of pragmatic competence. In this article, we argue that mastery of modal particles is important for interpersonal aspects of language development. The difficulty of learning and teaching German modal particles is outlined with regard to their theoretical explication. Investigating a group of L1 English university students, we show how a teaching program improved initially poor comprehension and production skills in the area of modal particles.

The authors of this study investigate the acquisition of German modal particles in tutored contexts. Such particles appear much more frequently in conversation than in written texts (cf. Rudolph, 1991), and their function is illocutionary and interpersonal rather than propositional. In very general terms, they indicate the speaker's attitude towards the utterance as well as the intended perception on the part of the hearer (cf. Helbig, 1977; Weydt, 1979; Bublitz, 1978; Thurmair, 1989). For example:

> Das wird **schon** wieder heilen.
> [That will **(schon)** again heal]
> Don't worry, that'll heal!
>
> (Bublitz, 1978, p. 226)

> Sie bleiben **doch** noch?
> [You stay **(doch)** a little longer]
> You are staying a little longer, (aren't you?)
>
> (König & Requardt, 1991, p. 71)

The ability to understand and use German modal particles correctly therefore is a good indication of pragmatic competence. However, modal particles are rarely used by learners of German as a foreign language. If they are employed at all, their use is very often inappropriate to the context. This reflects a lack of sensitivity to an important

© 1995, Martina Möllering and David Nunan

feature of German communication that might lead to misunderstandings and/or misinterpretations. How students of German as a foreign language employ these particles, and whether they can be effectively taught, are therefore important issues for language pedagogy. Accordingly, the study presented here was carried out to provide information on how focused instruction on German modal particles can be incorporated into a teaching program, and whether students' proficiency in the use of modal particles is raised as a result of taking part in such a program.

Supposedly universal phenomena of language, such as the realization of politeness through indirectness (Brown & Levinson, 1987) or conversational turn-taking (Sacks, Schegloff, & Jefferson, 1974, in Ervin-Tripp, 1987) have been challenged as being language- and culture-specific by researchers working in the area of cross-cultural pragmatics (cf. Thomas, 1983, 1984; Tannen, 1984; Wierzbicka, 1985; Wierzbicka, 1991). Recognition that speech acts might be realized differently in different cultures and subcultures has led to a large body of research that attempts to identify these differences (e.g. Blum-Kulka & Olshtain, 1984; Blum-Kulka, House, & Kasper, 1989; Kasper & Blum-Kulka, 1993). This research has practical significance, because differences in realizing speech acts linguistically may create problems in cross-cultural communication. As Tannen (1984) states, the lexicogrammar of an L2 is learned against the cultural background of L1. Therefore pragmatic strategies would be applied according to L1. The sort of failure that occurs is well illustrated by Thomas (1983). She labels it "pragmalinguistic failure," referring to different linguistic coding of speech acts that correspond in both cultures. Pragmalinguistic failure may lead to either negative assessment of a single speaker in one particular speech event or may lead to the stereotyping of an entire cultural group.

Rall (1981) reports that her Spanish L1 Mexican students inquired about the "directness" of German speech acts such as "asking" or "uttering of opinion." Nonnative speakers might perceive directness by looking at the syntactic mode of the encoding of a particular speech act without perceiving the modifications brought about by the use of modal particles (see also House & Kasper ,1981). The following example of "uttering of opinion" might illustrate this:

a. Es ist nicht einfach, dieses Problem zu lösen.
[It is not easy this problem to solve]
This problem is not easily solved.

b. Es ist **ja** nicht einfach, dieses Problem zu lösen.
[It is (**ja**) not easy this problem to solve]
This problem is not easily solved (as you know).

c. Es ist **doch** nicht einfach, dieses Problem zu lösen.
It is (**doch**) not easy this problem to solve]
(But you know that) this problem is not easily solved.

Whereas native speakers might perceive **(a)** as a turn in a discussion as quite

abrupt, **(b)** and **(c)** involve the illocutor's anticipated point of view. In **(b)** a shared opinion is assumed, while **(c)** expresses the wish to overcome a perceived difference of opinion (cf. Weydt, 1983). Modal particles function on the interpersonal level of discourse. They create what Schiffrin (1987) in her work on discourse markers in English calls "conversational cohesion;" in the case of *doch* and *ja* with reference to shared knowledge. Schiffrin assigns a similar function to the discourse marker "y'know" in English, namely to refer to meta-knowledge about what speaker and hearer share and to meta-knowledge about what is generally known.

Since the publication of Weydt's (1969, in Weydt, 1981) seminal work on French correspondences to German modal particles, a substantial body of literature on the subject has emerged. The term modal particle (*Modalpartikel*), however, is not employed consistently in these publications. In prescriptive grammars, they were considered superfluous, stylistically dubious "filler words" (*Füllwörter; Flickwörter*) and therefore to be avoided (cf. Busse, 1992). Other terms used in the literature are "flavouring words" (*Würzwörter*) (Paneth, 1981), "intentional particles" (*intentionale Partikeln*) (Rall, 1981) and "toning particles" (*Abtönungspartikeln*) (cf. Weydt, 1981; Husso, 1981; Vorderwülbecke, 1981). In a number of publications (cf. Dalmas, 1990, 1992), though, the term "particle" is used without further specification. This is not only a terminological issue but also a question of definition. Thus, Wierzbicka (1986) employs the general term particle in referring to modal particles in German as well as to focus particles in English (such as *very*) without making this distinction explicit, as Abraham (1991b) does for different types of particles in German. Muhr (1989) criticizes the ambiguity in classifying particles in pointing out declinability as a criterion.

A comprehensive definition is given by Thurmair (1989), who distinguishes between a "broad" morphological definition of the term particle as the class of non-declinables and a "narrower" definition as the class of non-declinables without prepositions, conjunctions, and adverbs, the latter definintion being the one employed in the literature quoted above. According to her definition, the morphologically defined word class "particle" can be divided into various subclasses. This further distinction is described as a functional one. As many particles occur in different functions, criteria such as position within the sentence would determine whether a particle occurs in "modal particle function." The following example using the particle *doch* might illustrate the point:

1. Es gab Möglichkeiten in Fülle, **doch** scheiterten sie...
   There were plenty of opportunities **but** they failed...

2. Komm **doch** morgen!
   (Come [doch] tomorrow)
   Why don't you come tomorrow!

In No. 1 *doch* occurs in the function of a sentence adverb, in No. 2 it has modal particle function.

The homonymy of modal particles with other particle functions leads to differences in classification. Thus, the number of particles which are referred to as modal particles in the relevant literature varies considerably. Following different theoretical premises, the teaching materials that have been devised discuss different particles as modal particles and propose different criteria for identifying a modal particle.

There is general agreement, however, on the fact that the lexical meaning of modal particles is "blurred" (Abraham, 1991a), "bleached" (Vorderwülbecke,1981), that they have "little denotative" meaning (Helbig, 1977). Helbig (ibid.) described them as "illocutive indicators," a position that has been revised and elaborated upon (cf. Abraham, 1991a & b, Thurmair, 1989, 1991). Thurmair (1989) posits that modal particles do not create but modify illocutionary force. They indicate the illocutors' common knowledge and the speaker's or listener's suppositions and expectations, they create cohesion with previous utterances, or mark the speaker's evaluation of the importance of the utterance. Thus, they appear mainly in colloquial German where spontaneous dialogues occur (cf. Rudolph, 1991), as well as in fiction, theater, and film dialogues where they can function as an "alibi" -simulating colloquial speech. Thurmair (1989) states that *ja*, *doch*, and *eben* also occur in academic texts, a fact for which Luchtenberg (1989) provides a plausible explanation in arguing that—amongst other functions they fulfill—modal particles serve to refer to different types of "common-sense knowledge." In particular, the modal particles *ja* and *doch* appear to accomplish this function:

1. Ich will morgen mal zum Arzt. - Morgen ist **doch** Mittwoch.
   I want to go to the doctor's tomorrow. - (But) tomorrow is
   Wednesday. (ibid., p. 662)

2. Seit Saussure ist **ja** der Systemcharakter der Sprache.....
   Since Saussure the character of language as a system is
   [ja].... (ibid., p. 667)

In No. 1 *doch* refers to knowledge about the fact that doctors' offices in Germany are normally closed on Wednesday afternoons. In No. 2 reference is made to specialist knowledge of a particular group, in this case linguists. Thus, amending a proposition in an academic text with a modal particle such as *ja*, *doch*, or *eben* evaluates the proposition as common knowledge and therefore indisputable.

Luchtenberg (1989) argues that the acquisition of modal particles that refer to common-sense knowledge is difficult because non-native speakers might not be familiar with the sociocultural context that forms the basis for common-sense knowledge. Therefore pragmalinguistic as well as sociopragmatic failure might occur. Furthermore, the non-native speaker is said to experience difficulties in expressing group solidarity. Steinmüller (1981) argues in a similar vein, stressing the importance for the non-native speaker to understand the agreement the native-speaker group has about the context. S/he should be enabled to refer to the same context of experience and evaluation and to express this reference linguistically.

Research findings (Husso, 1981; Steinmüller, 1981; Weydt, 1981) provide an ambiguous picture of the relationship between language acquisition in general and the acquisition of modal particles. There is agreement on a much lower frequency of use by non-native speakers that is hypothesized to be due to either lack of understanding of the important functions of modal particles because they do not occur in the learner's first language, lack of contact with native speakers, lack of instruction, or a combination of these factors. The fact that the theoretical explication has not yet reached a comprehensive and thus satisfactory level is one of the reasons why the teaching of German modal particles is considered problematic (cf. Götze, 1993; Busse, 1992) and different methodological approaches have been advanced. The timing and the progression of instruction on modal particles have been considered in a number of publicatons. Busse (1992) asserts that modal particles are too complex a phenomenon to be taught to second language learners before a very advanced state. He claims that language contrast and cognitive strategies are necessary methodology and should be employed at an advanced level, as the "intuitive" methodology used at beginner level is considered inappropriate. Vorderwülbecke (1981) proposed detailed criteria concerning the choice of particles with regard to progression. In contrast to Busse, he argues for an early introduction of modal particles if communicative competence and pragmatic adequacy are to be achieved and proposes focusing on the meaning of modal particles in context, explaining their force within a dialogue, and using intra-lingual contrasting, which might be achieved by paraphrasing.

A more student-centred approach is advanced by Rösler (1983), who argues that the situations in which communication is relevant for the learner can only be determined in view of the actual group of learners under consideration and cannot be deduced from a given progression of speech events; an argument which is also put forward by Tarone and Yule (1989) in a more general consideration of sociolinguistic competence in second language acquisition. There is evidence that instruction can improve this type of competence. Wildner-Bassett (1984) studied the use of gambits by adult L1 German learners of English. Analyzing videotaped data of interactional tasks in role-enactment form, she found that, regarding ratios of gambit quality and gambit quantity (among other variables she tested for), instruction had a positive effect. Billmyer (1990) reports that the production of and reply to compliments by learners of American English was influenced positively by formal instruction. She compared the productions of learners in social interaction with native speakers outside the classroom and found that the subjects who had received instruction performed significantly better than a control group on measures of frequency of norm-appropriate compliments, level of spontaneity, repertoire of adjectives used, and appropriateness of reply. Bouton (1994) investigated conversational implicature, the means of expressing a message indirectly, in the context of second language learning. His subjects were international college students in the United States. The longitudinal study shows that it takes non-native speakers of English a long time to interpret implicatures in a way that is not significantly different from native speakers' interpretations if there is no systematic teaching of implicature.

836

A number of studies have inquired into the perception and production of modal particles by second language learners. In an empirical study, Harden and Rösler (1981) investigated the perception of the use of modal particles and arrived at the following results (p. 73):

1. The listeners' language competence is significant for an evaluation of the use of particles.
2. The relationship between the use of particles and the expression of emotions seems to have been recognized.
3. The speaker's accent influences the assessment more than the use of particles.
4. Learners who received instruction in German as a foreign language do not seem to perceive the communicative value of particles as very high.

They proposed a more student-centred approach to teaching to include the expression of emotions in the communicative context of the classroom. Their study presents a product-oriented approach, it reflects a state of competence, but the nature of instruction of the subjects is not reported on.

Two qualitative studies—both belonging to the category of classroom-research studies (cf. Nunan, 1989, 1991, 1992)—report on instruction in the area of German modal particles. Paneth (1981) describes instruction of German modal particles to English L1 university students that draws on audio-lingual as well as translation methods. Students were presented with taped German texts containing modal particles and their English translations. They were also provided with a list of functions for each particle under discussion, which—together with the tapes—they used to group occurrences of particles according to appropriate translation into English. They then provided their own translations of the given and additional texts. Paneth (p. 108) observed that "in general the subjects find it easier to specify the meaning of a modal particle by translating rather than by choosing between alternatives or paraphrasing...." It seems difficult to determine to which extent this is a result of the teaching method as such because no detail is given on the language course as a whole. It seems contradictory, however, to use translation as a main means of instruction for a feature of language that is in a sense untranslatable. Whereas translation might facilitate comprehension of modal particles, it is difficult to imagine how it could lead to successful use in context.

Rall (1981) describes a teaching and learning project with Mexican university students after six semesters of instruction in German. Instruction started with a presentation of material in Spanish on modal particles in general and the problem of homonymy in particular followed by a detailed description of 25 particles in their modal and non-modal occurrences with examples. The explanations were illustrated by using audiotaped and written texts. Exercises included reading tasks aiming at correct intonation; drills, if they appeared necessary in the classroom context; paraphrasing, which is reported to have caused difficulties; gap exercises and open-ended role play; and au-

thentic communication with visiting native speakers of German. Production data showed that 15 of the 25 modal particles students had received instruction on were used with varying degrees of naturalness and appropriateness. Rall (1981) reported a number of avoidance strategies in the use of modal particles and concluded that interaction in German was crucial to the learning process. While she exclusively describes the focused instruction she provided, it seemed necessary for the purpose of the study presented here to investigate prior input on modal particles, as the students' levels of competence were thought to be heterogeneous at the onset of instruction.

The purpose of the study was to examine how instruction can influence acquisition of modal particles. It was conducted as classroom research (e.g., Nunan, 1991) and it was therefore important to assess how instruction on modal particles could be incorporated into a given teaching program. Instruction is here operationalized as a three-step process of awareness-raising, explanation, and practice. Acquisition is seen as a twofold process of perception/comprehension and production (for a discussion, see Nunan, 1995). Thus, the subjects' competence regarding modal particles before the onset of instruction had to be assessed. Competence was operationalized as the ability to perceive the illocutionary and emotive force of modal particles, the ability to use modal particles in context, and metalinguistic knowledge about modal particles. Also assessed were the input on modal particles prior to instruction and the subjects' motivation regarding the acquisition of modal particles.

In order to address the general issue, four detailed research questions are advanced:

1. Has the subjects' performance in using the modal particle about which they received explicit instruction improved?
2. Has their understanding of the illocutionary force of this modal particle improved?
3. Do the subjects perceive any improvement?
4. Which parts of the instruction did they find helpful?

Method

Subjects

The subjects chosen for the study were a group of 23 Australian undergraduate students of German who were taught by one of the researchers. The students had varying backgrounds in learning the language, some having commenced the program at beginners' level, others having studied German in high school. A number of students within this group had lived in a German-speaking country for varying amounts of time. All of the students were enrolled in a second-year (third-year for those starting at beginners' level) language course. In the beginners' class the curriculum was broadly communicative in orientation. Students were exposed to a large number of spoken dialogues (tape program and classroom input) with a relatively high frequency of modal

particles. A number of modal particles had been taken up for metalinguistic discussion within the beginners' class. In the following years the course is divided into four components: literary texts, topical texts, grammar, and conversation.

To provide details of instructed as well as uninstructed learning experiences, a questionnaire on biographical data was evaluated. The picture that emerged shows that 9 out of the 23 students had started the course at the beginners' level, the others after studying German at the secondary level. Students in both groups had spent varying amounts of time in a German-speaking country; two students in the post-secondary group have a German background, one, a Dutch background (seen as significant as Dutch is, like German, a language rich in modal particles). Those three latter students indicated that they spoke German (or Dutch respectively) as well as English at home. The following table illustrates the distribution:

TABLE 1
Biographical data

	Time in German-speaking country			
	3-12 months	**4-6 years**	**German (Dutch) background**	**Total**
Beginners' level students	1	1	-	9
Post-secondary students	3	2	2+1	14
	4	3	2+1	23

Hence, nearly half of the students participating in the course had had the opportunity for contact with native speakers of German outside an instructed context.

Materials

Pre-Instruction: Questionnaire

The subjects were presented with a questionnaire consisting of five sections:

Section 1 consisted of open-ended questions to elicit biographical data on the student's particular acquisition process (see Subjects).

Section 2 consisted of two versions of an audiotaped dialogue (published in Weydt et al. 1983); one containing modal particles (version A), the other without the particles (version B). In the

form of open-ended questions, subjects were asked whether they understood both versions, whether they perceived any differences between the two versions, and to name any differences they noticed. They were then presented with a written transcript of the two dialogue versions and asked to underline the differences.

Section 3 focused on the subjects' metalinguistic knowledge concerning modal particles and on their attitude towards using and learning them. Here the subjects were also asked which modal particles they were familiar with. Probes took the form of open-ended questions.

Section 4 elicited subjects' perception of the illocutionary and emotive force of modal particles. They were asked for an evaluation of the two dialogue versions according to a list of adjectives (e.g. "natural," "cold," "wooden"). They had to specify to which degree (0-6) each adjective described version A and B. This rating task was published in Weydt et al. (1983) where a native-speaker base line (n = 82) is given.

Section 5 consisted of a cloze task: two short gap exercises in the form of dialogues, in which the subjects were required to fill the gaps with appropriate modal particles out of a given number of choices. The gap exercises, published in Weydt et al (1983), were written on the basis of native speaker intuitions about the correct suppliance of modal particles in these specific contexts.

Pre-Instruction: Concordance Data

A concordance of texts the subjects had read and discussed in class was drawn up to provide examples of modal particle occurrences in context.

Instruction

The modal particle *doch* was chosen as focus of instruction for three main reasons:

1. it was underrepresented in the pre-test data on production
2. according to the concordance students had had ample input consisting of *doch* (the concordance showed *doch* to be the particle most often used in modal particle function [41out of 68 occurrences altogether] in the texts the students had read in the literature and topical text component of the course)
3. it may appear in a number of functions other than that of modal particle and in a variety of sentence types, and is therefore very complex to teach/learn.

840

Materials for instruction consisted of a handout with a detailed explanation of *doch*. It introduced different occurrences of *doch* in functions other than as modal particle as well as its overall modal particle function. This was further elaborated upon by specifying the illocutionary force it carries in different sentence types. Throughout the explanations, examples were provided that had been drawn from the corpus of texts used for the concordance. It was possible to keep the examples short, on the single-sentence level, without losing situational context, because the students had read the texts and were thus familiar with the context of situation as well as with the wider context. In cases where the concordance did not provide appropriate examples, examples were created to fit in with the students' experiential knowledge.

In the next step, the students were asked to go back to a dialogue text they had read and discussed in class some weeks earlier. The text reflected a debate between a father and his son and showed ample use of *doch*. Students were asked to discuss in pairs the occurrences of *doch* and to classify them according to the explanation they had been given earlier. Most of the students had considerable difficulty in doing so, and in the ensuing class discussion, they stated that they found the explanation very complex and that they needed more time to go over it again.

The last stage required students to use *doch* in the context of drill exercises that were amended by further explanation of the issues of word order and intonation. Opportunities for additional questions were given and error feedback provided.

Post-Instruction

The subjects were presented with a questionnaire consisting of four sections:

> ***Section 1*** described three situations the subject might be in. S/he was asked to respond to the situations, which provide an optional context for the use of modal particles.
> ***Section 2*** consisted of two authentic examples of the modal particle in question in a context the subjects were familiar with. In the form of an open-ended question, they were asked to explain the meaning with respect to the context.
> ***Section 3*** again tested for suppliance of the particle, but this time in an obligatory context. The subjects were presented with short gap exercises in the form of dialogues; they were required to fill the gaps with appropriate modal particles out of a given number of choices, including the one instruction was focused upon. The exercises, published in Weydt et al. (1983), are based on native speaker intuitions about the correct use of modal particles in these contexts.
> ***Section 4*** consisted of a closed questionnaire, where the students were asked to give ratings on their progress and their perceptions of the value of different aspects of the instructional process.

Procedures

The study was carried out in the form of classroom research (e.g., Nunan, 1991) as one of the researchers taught the topical text component of the second-year language course the subjects were enrolled in. Thus, the pre-test, instruction and post-test were incorporated into the regular teaching of this component over a period of five weeks. Instruction consisted of three periods of about thirty minutes each. Answering the pre-test-questionnaire can be seen as the first part of the instruction process, as it served to raise students' awareness of modal particles. The questionnaire was followed up with a short discussion about the questions and exercises within class.

Analysis

The data were subjected to descriptive and interpretative analysis. Quantification of data was employed for the purpose of illustration.

Results

Pre-Instruction Data

Most students were unable to detect the force of modal particles by just listening to the two dialogue versions. Having access to the written version as well enabled most of them to state at least one function fulfilled by modal particles in version A.

TABLE 2
Listening comprehension - Perception of differences between versions A and B

Differences between dialogue versions A and B	
no differences	6
differences but no specification	4
different style	6
different length/speed	4
differences in intonation	1
interpersonal differences	1
no answer	1

TABLE 3
Listening comprehension - Perception of differences in content

Differences between the statements made by version A and B	
no differences	12
misperception of content	3
differences (no clear specification)	4
differences on interpersonal level	1
no answer	3

TABLE 4
Reading comprehension - Description of modal particle functions

particles serve to:	
emphasise	5
create friendly atmosphere	3
make dialogue more natural	3
express colloquial speech	3
enforce sentence meaning	3

The evaluation of the illocutionary and emotive force of the two versions did not differ significantly from native speaker judgements as tendencies were perceived in the same way.

TABLE 5
Evaluation of illocutionary/emotive force

	Dialogue A		Dialogue B	
	subjects	NS	subjects	NS
natural	4.9	5.7	2.4	2.8
cold	1	1.7	3.3	3.3
warm	4.1	4.4	2.2	2.7
wooden	1.5	1.4	3.6	5
fluent	5.4	6	3.4	2.7
authentic	5	5.7	2.7	2.7
weak in establishing contact	0.9	2.4	3.1	4
friendly	5	5.7	2.8	3.7

The data on metalinguistic knowledge revealed that students had a vague notion of the function of modal particles and were generally unaware of the terminology used to classify them.

TABLE 6
Metalinguistic competence - Terminology

Terminology	
unknown	13
modal particle	1
particle	1
words for emphasis	3
stress words ("Betonungswörter")	2
others	3

They described them as providing emphasis (7) and expressiveness (6) and as a part of colloquial, spoken language (4).

With one exception, students expressed great interest in learning more about modal particles because of their importance for authentic spoken German. As the most common reasons, they stated that the particles would make the language sound more complete (8), that they wanted to learn how the language is really spoken (5), and that in

using modal particles they could display fluency (5).

Students were presented with two short gap exercises in the form of dialogues (Weydt et al., 1983) and were asked to fill the gaps with the appropriate modal particles out of a given number of choices. For some of the gaps various choices constituted an appropriate answer. The particles that were provided in the gap exercise were *denn* (43 times in 3 gaps), *einfach* (20 times in 3 gaps), *ja* (10 times in 2 gaps) and *doch* (29 times in 6 gaps), which can be calculated into the following rates of suppliance:

TABLE 7
Suppliance rate of selected modal particles in stage 1

denn	14.3	*ja*	5.0
einfach	6.7	*doch*	4.8

The suppliance rate of *denn* coincided with the students' self report data on their use of particles (Table 9). They seemed to have a working knowledge of *denn* which is bound to the sentence type "question." A particle to be seen as strongly underrepresented is *doch*. With *denn* it belongs to the group of most frequently used modal particles, nevertheless, its suppliance rate is significantly lower. In three gaps where it competed with *einfach*, the latter was the preferred choice. This does not mirror its frequency in native speech (Rudolph, 1991). In those cases where it was the only option or *ja* was the only alternative, the subjects responded with a large number of incorrect choices.

When asked about what they remembered learning about modal particles in their course of studies, there was a considerable difference between input (e.g., concordance data) and what the students actually took in.

TABLE 8
Input in instructed context

no input	3
not much input	7
can't remember	4
listing of MP	5

Some students gave further specifications as to the content of the explanations they had received. The fact that modal particles' meaning varied according to context and was hard to describe and translate was the most frequent comment (5), followed by the very general explanation that they served to emphasise (2).

When asked specifically which modal particles they used themselves, students named the following:

TABLE 9
Modal particles used by the students (self report data)

	frequency
denn	11
ja	11
doch	8
eigentlich	8
mal	7
schon	4
aber	1
eben	1

A number of other particles (5) were mentioned that do not occur in modal particle function, three of them by the student who speaks Dutch as well. The modal particles listed above were mentioned by 15 of the students. One of the German-background students declared that she used "all modal particles" without listing any. Five students wrote that they rarely used them, but did not specify which ones. One student stated that she did not use them at all, and one student did not answer.

The student who claimed to use significantly more modal particles was the Dutch speaker, who named ten (three of which cannot be classified as modal particles). Surprisingly, the second German-background speaker named only two. Only two of the students who had spent some time in a German-speaking country claimed to use slightly more different particles (5), whereas another student did not specify which particles but stated that he started using them after having spoken German for a number of hours.

When asked whether they had tried to learn how to use modal particles, the group was divided, with ten students answering positively and twelve negatively. Most of the students who had attended the beginners' class made an attempt at learning how to use them, with the exception of the two students in the group who had been in a German-speaking environment. Thus, the majority of those who had spent some time in a German-speaking context had not made a conscious effort to acquire them.

The corpus that was set up to create a concordance totaled just under 30,000 words. The modal particles considered here represent the "core" of modal particles referred to in the literature (cf. Thurmair, 1989; Muhr, 1989; Weydt et al., 1983).

846

TABLE 10
Concordance data

	occurrence as modal particle	occurrence as homonym	total occurrence
aber	1	171	172
auch	26	139	165
bloß	1	2	3
denn	13	26	39
doch	41	27	68
eben	9	2	11
*eigentlich**	3/31	28/--	31
einfach	5	1	6
etwa	- - -	9	9
halt	- - -	- - -	- - -
ja	34	54	88
mal	2	16	18
nur	1	92	93
ruhig	- - -	5	5
schon	15	64	79
sowieso	2	- - -	2
vielleicht	- - -	23	23
wohl	15	3	18

* The differing figures reflect differing classifications: According to Thurmair (1989), only 3 occurences can be classified as modal particles, whereas the other 28 are said to belong to the category of sentence adverb. Weydt et al. (1983) employ a wider framework according to which all occurrences can be counted as modal particles ("Abtönungspartikeln").

With regard to different learner profiles, no discernible pattern emerged, but some students stood out in different areas. Hence, the student who had been in a German-speaking country longest was able to point out the force of modal particles on the interpersonal level of language, the Dutch-background student pointed out considerably more particles than other students, and the German-background students performed significantly better in the production exercises.

Post-Instruction Data

Students were provided with two examples of *doch* in use in contexts they were familiar with. These examples were taken from texts that had been discussed in class, but

847

without any focus on the use of *doch*. Nearly half of the students (the second questionnaire was answered by 22 subjects only) were able to either explain or translate the force of *doch* correctly in the two examples they were presented with, while about a third provided only a superficial explanation.

TABLE 11
Comprehension in Stage 2 - Explanation of doch

	example 1	example 2
correct explanation	8	7
vague explanation	7	7
incorrect explanation	2	3
correct translation	2	1
incorrect translation	3	4

The production data showed that *doch* was infrequently supplied in non-obligatory contexts, though with more frequency in a situation and sentence type (Situation 3) that had been part of the drill exercises discussed in class.

TABLE 12
Production data in stage 2 (suppliance in non-obligatory context)

	doch correct	*doch* incorrect	no suppliance
situation 1	4	3	15
situation 2	4	6	12
situation 3	8	2	12

The gap exercises yielded a much higher correct suppliance rate of *doch* than before instruction had taken place. Whereas it was 4.5 before instruction, it rose to 10.5 thereafter, *doch* is suggested also as an incorrect option rather frequently (4.5), an occurrence that can be interpreted as overgeneralization (see discussion).

Overall, though, the students' incorrect productions in the gap exercises seem to have risen. Taking into consideration that before instruction they were presented with 9 gaps (compared to 11 gaps after instruction), the rate of inappropriate suppliance rose from 3.9 to 4.5. None of the students supplied fewer than three inappropriate options, the students with German backgrounds not faring much better (3 and 6 inappropriate choices) than the NNS subjects. More than 30% of incorrect choices, how-

ever, were made in two gaps. The correct modal particles to be supplied here were either not mentioned at all or very infrequently in the students' self-report data on use and did not occur at all or very infrequently in the concordance data. If these two gaps are not considered, the rate for 9 gaps has fallen from 3.9 before instruction to 3.4 after instruction—a small improvement overall.

The subjects' evaluation also reflected this tendency to overall improvement and judged the explanation with examples as the most effective means of instruction.

TABLE 13
Student evaluation of instruction

Question	Rating (out of 5)
1. improvement in understanding *doch*	3.3
2. improvement in using *doch*	3.1
3. improvement in awareness about MP in general	3.2
4. effectiveness of explication	3.6
5. effectiveness of examples	3.7
6. effectiveness of drill exercises	3.4

Summary

Before instruction, the group of students under study showed poor comprehension of the force of modal particles. Their metalinguistic knowledge was restricted to a vague notion of the functions modal particles fulfill and they reported to have had little or no input on the topic in their course of studies. Input in the form of occurrences of modal particles in context (concordance data) without focus on the particles seems to have had no influence on the subjects' performance, as the underrepresentation of *doch* reflects. There was little difference in their evaluation of the emotive and illocutionary force of modal particles compared to native-speaker judgements. Self-report data indicated that students used only a small number of modal particles with *denn, ja, doch* and *eigentlich* being mentioned by at least a third of the students. Their use of modal particles in a given context deviated from native-speaker expectations, especially with regard to the particle *doch*. Production data revealed an underrepresentation of *doch* although it had occurred in modal particle function more often than any other particle in the texts the students had read, as the concordance data showed.

Thus, *doch* was chosen as the focus of instruction and was used with considerably more frequency after the instruction phase, though in what might be called "semi-obligatory" contexts rather than optional ones. Less than half of the students were able to explain or translate *doch* in a way which showed that they fully understood its illocutionary force. Thus, an improved understanding cannot be posited for the

group as a whole but should be seen as a tendency. Acquisition is a complex "organic" rather than linear "all-or-nothing" process. Instruction is therefore interpreted as having had an influence on the acquisition of *doch*, though some aspects of the validity of this conclusion require further consideration (see below).

The students showed great interest in the acquisition of modal particles and they perceived the instruction as having had a positive effect. They found the detailed explanations and the accompanying examples the most effective means of instruction, but remarked that they needed more time to grasp the distribution of meaning of *doch* according to function and sentence type. Except for the significantly better performance of German-background speakers before instruction, no pattern of differences could be found between learners of different backgrounds.

Overall, the instruction is seen to have had positive effects. There are, however, some threats to the reliability and validity of the study, which will be set out below.

Discussion

The listening comprehension data on differences between the two dialogue versions corroborated the findings by Harden and Rösler (1981) that learners who received instruction in German as a foreign language did not seem to perceive the communicative value of particles as very high, as only a third of the students perceived the function of modal particles in dialogue A. It should be considered, however, that the number of subjects is rather small and that external validity might therefore be limited. In evaluating illocutionary and emotive force, the subjects (n = 23) have been compared to a much larger group of native speakers (n = 82) which could restrict the reliability of the findings.

One thing we found was that suppliance of *doch* varied considerably according to context. This confirms the observation of Kasper and Dahl (1991) in their analysis of elicitation methods. Husso (1981) also pointed out the limited validity of her findings with regard to elicitation, which was—as is the case in this study—a gap exercise. A change in the suppliance of one particular item in the same type of elicitation procedure, though, is to be seen as a valid indication of the influence of instruction. We would not suggest that this proves acquisition of *doch* has occurred, as it is not frequently employed in optional contexts, which might be due to avoidance strategies (Færch & Kasper, 1983) or to the ambiguity of the situation described. The latter seems to hold true especially for the first "simulation" the students were asked to undertake. It seems that the situation was not defined clearly enough for them to perceive an optional context for the use of *doch*. A similar problem was identified when using the gap exercises. The exercise book they were taken from (Weydt et al., 1983) proposes a certain set of correct options per gap. The students made other choices, however, that could be seen as correct when interpreting the context of situation in a slightly different way, although the particles they proposed were not always in modal particle function. We would therefore agree with Muhr (1989) and Götze (1993), who argue for a very detailed description of the context of use.

850

We would qualify the provision of *doch* as incorrect choice after instruction had taken place as overgeneralization of the possible contexts of use, similar to the effects Lightbown (1983), Pica (1985), and Brindley (1987) observed in their studies on morphosyntactic acquisition. As *doch* had been the particle focused upon, it was the most likely one to be supplied when in doubt. The latter tendency as well as the fact that only short-term effects are indicated, limits the validity of the study.

The frequent use of the modal particle *denn* reported by Husso (1981) was confirmed by the present study. Vorderwülbecke's (1981) progression criteria provide a plausible explanation. *Denn* belongs to the group of most frequently used modal particles and it is bound to the sentence type "question," which considerably limits the contexts of its use and thus renders it less complex to learn than *doch*.

The time set aside for instruction on *doch* was rather limited. More time would have been necessary to ensure that most of the students had fully grasped the explanations provided. More practice would have been beneficial, especially with respect to comprehending and explaining the illocutionary force of *doch* in different contexts. The limited time frame made it impossible to discuss combinatorial and intonational aspects of *doch* in great detail.

The study is seen as providing evidence for the beneficial effects of instruction on the acquisition of modal particles although we must reiterate that the modal particle focussed upon—*doch*—was not fully acquired in all its complexity. We believe that there should be more emphasis on modal particles in the instructional process. The explanations provided should increase in complexity according to the students' level of competence and should, at the latest stage, not only enable the student to use modal particles correctly but also to be able to explain their illocutionary and emotive force in context. This seems particularly important for those university students envisaging a career in teaching, who would need to be able to communicate about them as well as use them.

A number of spoken texts in the form of taped interviews and conversations of the debate style described above were used in the language course. The concordancing procedure showed them to contain a large number of modal particles. These texts lend themselves to raising awareness of the use of modal particles and to pointing out and discussing their use in different contexts. Using spoken texts rather than written material provides the necessary opportunities to increase students' listening competence. An important subskill here is the ability to detect the illocutionary force of a spoken text. This seems again particularly important with regard to aspiring language teachers (cf. Rösler,1983; Luchtenberg, 1989). As Frescura (1991) put it in a proposal for more focus on listening comprehension in university language courses:

> [L]anguage teachers tend to teach what they have been taught. As students, they had grammar classes, conversation classes but not "understanding the spoken word classes." (p. 119)

Additionally, the explanation provided with the help of examples collected through the concordancing procedure for *doch* may be taken as a model for explanation of other modal particles, which can then be used as part of a core language course. Overall the spoken language should receive more attention, a conclusion that seems justified by the students' interest in authentic spoken language.

References

Abraham, W. (1991a). Discourse particles in German: How does their illocutionary force come about? In W. Abraham (Ed.), *Discourse particles* (pp. 203-252). Amsterdam: John Benjamins.

Abraham, W. (1991b). Modal particle research: The state of the art. *Multilingua, 10*(1&2), 9-15.

Billmyer, K. (1990). "I really like your lifestyle:" ESL learners learning how to complement. *Pennsylvania Working Papers in Educational Linguistics, 6*(2), 31-48.

Blum-Kulka, S., & Olshtain, E. (1984). A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics, 5*(3), 196-213.

Blum-Kulka, S., House, J., & Kasper, G. (Eds.), (1989). *Cross-cultural pragmatics: Requests and apologies.* Norwood, NJ: Ablex.

Bouton, L.F. (1994). Conversational implicature in a second language: Learned slowly when not deliberately taught. *Journal of Pragmatics, 22*, 157-167.

Brindley, G. (1987). Verb tenses and TESOL. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 173-204). Adelaide, Australia: National Curriculum Resource Centre.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language use.* Cambridge. Cambridge University Press.

Bublitz, W. (1978). *Ausdrucksweisen der Sprechereinstellung*. Tübingen: Niemeyer.

Busse, D. (1992). Partikeln im Unterricht Deutsch als Fremdsprache. *Muttersprache, 102*, 37-59.

Dalmas, M. (1990). Partikelforschung 'konkret.' *Deutsch als Fremdsprache, 27*(5), 285-289.

Dalmas, M. (1992). Bewertung durch Partikeln. *Deutsch als Fremdsprache, 29*(2), 90-95.

Ervin-Tripp, S. (1987). Cross-cultural and developmental sources of pragmatic generalizations. In J. Verschueren & M. Bertuccelli-Papi (Eds.), *The pragmatic perspective* (pp. 47-60). Amsterdam: John Benjamins.

Færch, C., & Kasper, G. (1983). Strategies in interlanguage communication.

Götze, L. (1993). Modalpartikeln aus der Sicht der Sprachlehr- und Sprachlern forschung. *Deutsch als Fremdsprache, 30*(4), 227-232.

Harden, T. & Rösler, D. (1981). Partikeln und Emotionen - zwei vernachlässigte Aspekte des gesteuerten Fremdsprachenerwerbs. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp. 67-80). Heidelberg: Groos.

Deutschunterricht (pp. 67-80). Heidelberg: Groos.

Helbig, G. (1977). Partikeln als illokutive Indikatoren im Dialog. *Deutsch als Fremdsprache, 14*, 30-44.

House, J., & Kasper, G. (1981). Politeness markers in English and German. In F. Coulmas, *Conversational Routine*. The Hague: Mouton.

Husso, A. (1981). Zum Gebrauch von Abtönungspartikeln bei Ausländern. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp.81-99). Heidelberg: Groos.

Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition, 13*(2), 215-247.

Kasper, G., & Blum-Kulka, S. (Eds.), (1993). *Interlanguage pragmatics*. New York: Oxford University Press.

König, E., & Requardt, S. (1991). A relevance-theoretic approach to the analysis of modal particles in German. *Multilingua, 10*(1&2), 63-77.

Lightbown, P. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. Long (Eds.), *Classroom- oriented research in second language acquisition* (pp. 217-245). Rowley, MA: Newbury House.

Luchtenberg, S. (1989). Partikeln und Alltagswissen. In H. Weydt (Ed.), *Sprechen mit Partikeln* (pp. 661-672). Berlin: de Gruyter.

Muhr, R. (1989). Zur Didaktik der Modalpartikel im Unterricht Deutsch als Fremdsprache. In H. Weydt (Ed.), *Sprechen mit Partikeln* (pp. 645-660). Berlin: de Gruyter.

Nunan, D. (1989). *Understanding language classrooms*. London: Prentice Hall.

Nunan, D. (1991). Methods in second language classroom-oriented research. A critical review. *Studies in Second Language Acquisition, 13*, 249 -274.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Nunan, D. (1995). Closing the gap between learning and instruction. *TESOL Quarterly, 29*(1), 133 - 158.

Paneth, E. (1981). Partikeln im Unterricht—Erfahrungen mit englischen Studenten. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp. 101-110). Heidelberg: Groos.

Pica, T. (1985). The selective impact of classroom instruction on second language acquisition. *Applied Linguistics, 6*(3), 214-222.

Rall, M. (1981). 'Se puede ensenar la necesidad de emplear particulas intencionales?' Ein Experiment mit spanischen Studenten. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp. 123-136). Heidelberg: Groos.

Rösler, D. (1983). Der Erwerb der Abtönungspartikeln im institutionalisierten Lernprozeß Deutsch als Fremdsprache. In H. Weydt (Ed.), *Partikeln und Interaktion* (pp. 291-300). Tübingen: Niemeyer.

Rudolph, E. (1991). Relationships between particle occurrence and text types. *Multilingua, 10*(1&2), 203-219.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Steinmüller, U. (1981). Akzeptabilität und Verständlichkeit—Zum Partikelgebrauch von

Heidelberg: Groos.
Tannen, D. (1984). The pragmatics of cross-cultural communication. *Applied Linguistics, 5(3)*, 189-195.
Tarone, E., & Yule, G. (1989). *Focus on the language learner*. Oxford: Oxford University Press.
Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4*(2), 91-112.
Thomas, J. (1984). Cross-cultural discourse as 'unequal encounter': Towards a pragmatic analysis. *Applied Linguistics, 5(3)*, 226-235.
Thurmair, M. (1989). *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer.
Thurmair, M. (1991). 'Kombinieren Sie doch nur ruhig auch mal Modalpartikeln!': Combinatorial regularities for modal particles and their use as an instrument of analysis. *Multilingua, 10*(1&2), 19-42.
Vorderwülbecke, K. (1981). Progression, Semantisierung und Übungsformen der Abtönungspartikeln im Unterricht Deutsch als Fremdsprache. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp. 149-160). Heidelberg: Groos.
Weydt, H. (1979). *Die Partikeln der deutschen Sprache*. Berlin: de Gruyter.
Weydt, H. (1981). Methoden und Fragestellungen der Partikelforschung. In H. Weydt (Ed.), *Partikeln und Deutschunterricht* (pp. 45-64). Heidelberg: Groos.
Weydt, H. (1983). *Partikeln und Interaktion*. Tübingen: Niemeyer.
Weydt, H. (1989). *Sprechen mit Partikeln*. Berlin: de Gruyter.
Weydt, H., Harden, Th., Hentschel, E., & Rösler, D., (1983). *Kleine deutsche Partikellehre*. Stuttgart: Klett.
Wierzbicka, A. (1985). Different cultures, different languages, different speech acts. Polish vs English. *Journal of Pragmatics, 9*(2&3), 145-178.
Wierzbicka, A. (1986). Introduction. (Special issue on particles ed. by A. Wierzbicka). *Journal of Pragmatics, 10*, 519-534.
Wierzbicka A. (1991). *Cross-cultural pragmatics: The semantics of human interaction*. Berlin/New York: Mouton de Gruyter.
Wildner-Bassett, M.E. (1984). *Improving pragmatic aspects of learners' interlanguage*. Tübingen: Narr.

854

Authors

MARTINA MÖLLERING, Lecturer in German Studies, School of Modern Languages, Macquarie University, Sydney, New South Wales, 2109, Australia. Specializations: pragmatic aspects of second language acquisition, computer-assisted language learning.

DAVID NUNAN, Professor, English Centre, The University of Hong Kong, Pokfulam Road, Hong Kong. Specializations: curriculum development, learning strategies, language teaching methodology.

Applied Language Learning
1995, Vol. 6, Nos. 1 & 2, pp. 65-87

The Sociolinguistic Patterns of Native Arabic Speakers: Implications for Teaching Arabic as a Foreign Language

Anwar A. Hussein
Ottawa University

This article presents a brief descriptive analysis of speech acts in Arabic: forms of address, apologies, requests, expressions of gratitude, disagreement, greetings, refusals, partings, and telephone etiquette. A contrast with American speech acts is also presented in the process of description. The author gathered data by means of participant observation in daily situations and then analyzed them to disclose their usage and function. Results indicated that linguistic formulas of each speech act were determined by social distance, formality of the situation, age, level of education, and status of the participants. Moreover, the contrast between Arabic and English speech acts revealed that there are fundamental differences between the sociolinguistic patterns of Arabic and American native speakers, not only in verbal communication but also in nonverbal communication. The information dwells on the nature of communicative competence, first and second language learning and acquisition, and it is hoped that it will ease intercultural communication between Arabic and American native speakers.

The field of teaching Arabic as a foreign language (TAFL) or as a second language (TASL) is a new one, and it is growing rapidly in the United States, in Europe, and throughout the world. Learners of Arabic, particularly those who master it for specifically communicative purposes, need to be competent not only linguistically, but also socially. That is, they need to know how to use the language with its native speakers according to native social rules and cultural norms. Culture has a considerable impact not only on our values, beliefs, and social interactions, but also on how we view the world, what we consider important, what we attend to, and how we learn and interpret information (Philips, 1983; Delgado-Gaitan & Trueba, 1985; Huitt, 1988; Jacobs & Jacobs 1988; Jacobs, 1990; Rhodes, 1990).

© 1995, Anwar A. Hussein

To acquire a native or near-native command of the target language, the learner must attain a high competency in the three components of communication that Saville-Troike (1989) delineated. These are: (a) linguistic knowledge (e.g., verbal and nonverbal elements, syntax, and semantics), (b) interaction skills (e.g., rules of speech, discourse organization and processes, and norms of interaction and interpretation), and (c) cultural knowledge (e.g., social structure, values and beliefs, attitudes, and schemata). Foreign-language learners must gain cultural competence (i.e., the ability to function comfortably and successfully in cross-cultural settings and to interact harmoniously with people from other cultures that differ from their own). The research on effective cultural communication indicates that culturally competent individuals (a) cope effectively with the psychological and emotional stress of dealing with the unfamiliar, (b) quickly establish rapport with others, (c) sense other people's feelings, (d) communicate effectively with people from varying backgrounds, and (e) respond adequately to miscommunication (Giles, Coupland, Williams, & Leets, 1991). In addition, effective cross-cultural communication requires knowledge of nonverbal communication. Nonverbal cues set the stage for interpersonal communication (Barlund, 1968; Hall, 1973; Curt, 1976; LaFrance & Mayo, 1978) and "are critical components of participants' messages" (Bonivillain, 1993, p. 37).

The purpose of this article is to discuss some of the sociolinguistic patterns of Arabic speakers pertaining to forms of address, apologies, requests, expressions of gratitude, greetings, refusals, partings (leave-takings), disagreements and telephone etiquette. Such brief delineation of the aforementioned speech acts is intended to help non-native Arabic speakers avoid appearing impolite, unfriendly or boorish, as Thomas (1983, p. 96-97) pointed out in her comment on the linguistic pattern of foreign-language learners:

> If a non native speaker appears to speak fluently (i.e., grammatically competent), a native speaker is likely to attribute his/her apparent impoliteness or unfriendliness not to any linguistic deficiency, but to boorishness or ill-will. While grammatical error may reveal a speaker to be a less than proficient language user, pragmatic failure reflects badly on him/her as a person.

Methodology

The author collected data as a participant in and as an observer of real-life situations (celebrations, meetings, social visits at homes, informal gatherings) among native Arabic speakers and analyzing written discourse in newspapers and letters. Every effort was made to collect data from as broad a range of occupational and educational backgrounds as possible. However, most of the data were collected from graduate students and professional people (e.g., university and college instructors, medical doctors, and business people). Moreover, the reader should note that the analyses of

857

speech acts presented here do not represent all speakers of the Arabic language. In particular, these speech acts represent the sociolinguistic pattern of Muslim educated Palestinian and Jordanian speech communities. Thus, the work reported here is the result of a naturalistic inquiry approach to data collection and analysis.

Conceptual Framework

Achieving comfortable and successful intercultural communication requires not only fluency in the target language, but also knowledge pertaining to the rules governing speech acts in the target culture. Landis, Brislin, and Hulgus (1985), Barna (1983), and Oberg (1986) found that anxiety caused by intercultural interaction is one of the barriers to the creation of comfortable and successful crosscultural communication. C. Stephan (1992), and the team of C. Stephan and W. Stephan (1985, 1989a, 1989b) hypothesized that anxiety in intergroup interactions is the result of four categories of feared negative consequences. These are: (a) negative psychological consequences (e.g., frustration or loss of control), (b) negative behavioral consequences (e.g., exploitation or verbal derogation), (c) negative evaluations by outgroup members (e.g., stereotyping or disdain), and (d) negative evaluations by in-group members (e.g., disapproval or rejection for having contact with the outgroup).

Moreover, intergroup anxiety is the result of three main variables. These are: (a) prior intergroup relations, (b) situational factors, and (c) prior intergroup cognitions. The amount and conditions of contact, status differences between group members, and attitudes of significant others towards the outgroup are some of the significant factors of prior intergroup relations. Some of the situational factors are degree of structure, type of interdependence and group composition, stereotyping and prejudice. Ethnocentrism and knowledge of the culture of the outgroup are some of the prior intergroup cognition factors (Stephan, 1992; Stephan & Stephan,1985, 1989a, 1989b).

Ignorance of a group has been associated with negative, unsuccessful, and uncomfortable intergroup relations (Davidson & Thomson, 1980). Being aware of and familiar with the target culture is associated with positive, accurate, and comfortable social interactions (Berry, 1970; Brewer & Campbell, 1976). Ignorance or insufficient information about the target culture leads to stereotypes that occupy a significant role in intergroup interactions. Stereotypes against the target culture underlie expectations, and they influence perceptions, behaviors, and effective interactions with the native people of the target culture (Neuberg, 1989).

Some of Sociolinguistic Behavior of Native Arabic Speakers

Forms of Address

Forms of address are vital for foreign language learners who intend to socialize with and live among the native speakers of the target language. Commenting on the significance of forms of address, Philipsen and Huspeck (1985, p. 94) state that

speech acts presented here do not represent all speakers of the Arabic language. In particular, these speech acts represent the sociolinguistic pattern of Muslim educated Palestinian and Jordanian speech communities. Thus, the work reported here is the result of a naturalistic inquiry approach to data collection and analysis.

Conceptual Framework

Achieving comfortable and successful intercultural communication requires not only fluency in the target language, but also knowledge pertaining to the rules governing speech acts in the target culture. Landis, Brislin, and Hulgus (1985), Barna (1983), and Oberg (1986) found that anxiety caused by intercultural interaction is one of the barriers to the creation of comfortable and successful crosscultural communication. C. Stephan (1992), and the team of C. Stephan and W. Stephan (1985, 1989a, 1989b) hypothesized that anxiety in intergroup interactions is the result of four categories of feared negative consequences. These are: (a) negative psychological consequences (e.g., frustration or loss of control), (b) negative behavioral consequences (e.g., exploitation or verbal derogation), (c) negative evaluations by outgroup members (e.g., stereotyping or disdain), and (d) negative evaluations by in-group members (e.g., disapproval or rejection for having contact with the outgroup).

Moreover, intergroup anxiety is the result of three main variables. These are: (a) prior intergroup relations, (b) situational factors, and (c) prior intergroup cognitions. The amount and conditions of contact, status differences between group members, and attitudes of significant others towards the outgroup are some of the significant factors of prior intergroup relations. Some of the situational factors are degree of structure, type of interdependence and group composition, stereotyping and prejudice. Ethnocentrism and knowledge of the culture of the outgroup are some of the prior intergroup cognition factors (Stephan, 1992; Stephan & Stephan,1985, 1989a, 1989b).

Ignorance of a group has been associated with negative, unsuccessful, and uncomfortable intergroup relations (Davidson & Thomson, 1980). Being aware of and familiar with the target culture is associated with positive, accurate, and comfortable social interactions (Berry, 1970; Brewer & Campbell, 1976). Ignorance or insufficient information about the target culture leads to stereotypes that occupy a significant role in intergroup interactions. Stereotypes against the target culture underlie expectations, and they influence perceptions, behaviors, and effective interactions with the native people of the target culture (Neuberg, 1989).

Some of Sociolinguistic Behavior of Native Arabic Speakers

Forms of Address

Forms of address are vital for foreign language learners who intend to socialize with and live among the native speakers of the target language. Commenting on the significance of forms of address, Philipsen and Huspeck (1985, p. 94) state that

point that non-native speakers must be aware of, is that it is not appropriate to call a married woman who has children by her first name. The form of address *yaa um* + the name of her oldest son must be used. If a woman has no sons, she can be called by her oldest daughter's name. But if she has no children, usually people call her by the name that she intends to name her first son. When in doubt, a speaker can address a female as *yaa ukht/ukhtii*, *yaa aanisah*, *yaa saiyydah*. It is more appropriate and respectful to use the address form *yaa ukht* or *yaa ukhtii* in such situations. However, in informal situations, close friends and family members might use the first name of a married friend or daughter. The same can be applied in addressing men, but using the masculine forms. The following forms are used in addressing men.

TABLE 2
Forms of address for men

Form of address	Gloss	Used with	Style
saiyyid	Mr.	Nonacquaintances	Formal
akh\akhii	Brother/My brother	Nonacquaintances	Formal
First name	————	Friends	Informal
abuu (e.g., *abuu Ahmed*)	Father of	Friends and Acquaintances	Formal/Informal
khalii + 1st Name	My uncle	One's mother's brother	Informal
ʿamii + 1st Name	My uncle	One's father's brother	Informal
ʿami	Father-in-law	One's husband's father	Informal
hajjii	————	All interlocutors	Formal/Informal

It is worth mentioning that kinship terms (See Table 3) are used also as polite terms for nonrelatives.

860

TABLE 3
Kinship terms used with nonrelatives

Kinship term	Gloss
ʿam	uncle (father's brother)*
ʿamah	aunt (father's sister)*
jad	grandfather
jadah	grandmother
ibn alʿam	cousin (male)
ibnat alʿam	cousin (female)
khaal	uncle (mother's brother)*
khaalah	aunt (mother's sister)*

* Note: The only significance behind using these different terms is that the addressee is from the father or mother's family side.

Another aspect of address etiquette that must be taught to non-native speakers is how titles are used with names. In contrast to English, in Arabic the first name or both the first name and the second (middle) name are used with the specific title, which depends on the interlocutor's status and position. First and second names plus title are used in writing (e.g., letters and memos) and in the formal introduction of a speaker or a guest to other people. The following examples illustrate this usage:

1.	*jalaalat almalik* Hussein bin Talal	(*jalaalat almalik*: the majestic king)
2.	*sumuw alamiir* Abdallah bin Abedalcaziz	(*sumuw alamiir*: his highness)
3.	*sumuw alamiirah* Hanan Alhussein	(*sumuw alamiirah*: her highness)
4.	*fakhaamat/siyyaadat alraiis* Ali Abdallah Saleh	(*fakhaamat*: greatness) (*siyyaadat*: Mastery of)
5.	*maʿaalii raiis alwuzaraa?* Zeid Alrifaʿi	(*maʿaalii*: highness)
6.	*mudiir altarbiyah waltaʿliim* Zahran Hasunah	(*mudiir*: superintendent)
7.	*mudiirat altarbiyah waltaʿliim* Lamah Hashim Saleh	(*mudiirat*: superintendent)
8.	*alsheikh* Ahmad Yasin	(*alsheikh*: the sheikh)
9.	*alsheikhah* Zeinab Taha	(*alskeikhah*: the sheikh)
10.	*alimaam* Ahmed Saqir	(*alimaam*: The person who leads the prayer)
11.	*alkhaitiib* Jamal ʿatiyah	(*alkhatiib*:The person who delivers the Friday speech prayer)*

861

12. *alduktor* Anwar Hussein	(*alduktor*: the doctor)
13. *alustaath* Sameh Saleh	(*alustaath*: the teacher)
14. *almuʿalimah* Reem Hashim Saleh	(*almuʿalimah* [feminine]: the teacher)
15. *ṭabiib* (No names are used with this title)	(*ṭabib*: medical doctor)
16. *ṭabiibah*	(*ṭabiibah* [feminine] medical doctor)
17. *alḥakiim*	(*alḥakiim*: medical doctor)
18. *ḥadrit alakh* Samir Al-Ghazali	(*ḥadrit*: his honor)
19. *ḥadratikum*	(your honor)

* Note: Each Friday at 12:30 p.m., there is a speech that lasts approximately 30 minutes before the second prayer.

From the above examples we notice that with almost all the titles, the first name and the family name are used. Sometimes the word *bin* "son of" is inserted between the first and the family name or the second (middle) name. Family names are never used alone to refer to any person. The titles *ṭabiib* and *ṭabiibah* refer to medical doctors, but native Arabic speakers rarely use these standard forms in their colloquial speech. Instead, they prefer to use the terms *alduktor* (referring to an M.D. or a Ph.D.) or *alḥakiim*, which is equivalent to *alduktor* (referring only to an M.D.) because they sound more prestigious, especially the former form. The last title, *ḥadrit*, is used before all the other titles at the beginning of formal written letters. Note these examples:

ḥadrit maʿaalii alwaziir ṭaahir almaṣrii
(His honor and highness the minister Tahir Almasri.)

ḥadrit mudiir altarbiyah waltaʿliim Ahmed Saleh
(His honor the superintendent Ahmed Saleh)

The plural form *ḥadratikum* is used when requesting someone's presence or a favor, even if there is only one addressee. Note the following examples (a wedding invitation and a request for help respectively):

natasharrafu bidaʿwati ḥadratikum ilaa ḥoduur ḥaflat zafaaf waladaynaa . . .
(We have the honor of inviting you to attend the wedding of our son and daughter . . .)
narjuu min ḥadrtikum an tabthuluu fii maa wasʿikum limusaʿadat al?akh ...
(We beg that your honor do your best in helping the brother . . .)

The use of the plural is another way of showing respect.

Learners of Arabic should be aware of the Arabic address system, and particularly of the pragmatic use of those address forms that are employed extensively in social situations (See Tables 1 and 2) to achieve smooth and effective communication. Not being aware of the appropriate use of such forms leads to miscommunication and to unintentional insults. Using the "no naming" strategy (Brown & Ford, 1961) common among native American English speakers does not work when dealing with Arabs; rather it compounds the problem because forms of address and names must be used to indicate status and respect for the interlocutors. Such signals are especially important when a person addresses someone of a higher status. Janicki (1985, p. 44) states on the negative attitudes that might result from misusing the address system in a particular culture,

> A mishandling of the address system may be taken usually to generate negative attitude, but the dimension congenial, uncongenial is most likely to notify that reaction. If the foreigner is assigned a high estimate on that dimension, the native speaker's potentially negative attitude may turn into a positive one. If, however, the foreigner is assigned a low estimate on the dimension, then the attitudes takes on a negative value.

Apologies

In any society, linguistic formulas of apology, nonlinguistic means, or both serve to maintain what Goffman (1971) called "the public order." Speakers use such means to remedy potentially unpleasant social situations in order to maintain smooth social interaction. In studying this speech act in American English, Owen (1980) organized strategies for initiating a repair into three categories. These categories are (a) nonsubstantive (ritual) strategies, (b) semisubstantive strategies, and (c) substantive strategies. The ritual category includes asserting imbalance or showing deference, asserting that an offense has occurred, expressing an attitude toward an offense, and requesting the restoration of imbalance. Giving an account falls under the second category of semisubstantive strategy. Giving an account was put into an independent category because, as Wolfson (1990) found, native speakers of American English do not equate an account with an actual apology. The substantive category includes such nonlinguistic strategies as repairing the damage and providing compensation. "The extent to which either ritual or substantive remedies are considered appropriate in a particular case will depend on features of the offense and on cultural criteria" (Owen, 1980, p. 170).

In Arabic culture, even though an offended person is permitted to demand compensation or repair of the damage, that person is strongly encouraged to forgive and say *Allaah bi'awidnii,* "Allah (God) will compensate me for that," particularly if the damage is not severe and was caused inadvertently. In any event, the person who

caused the offense (regardless of its effect, personal damage, or property damage) must apologize and ask for forgiveness. The three main key words in linguistic formulas of apology in Arabic language are *aasif*, "Sorry"; *ʿafwan*, "I seek your forgiveness;" and *maʿthiratan*, "My apology." The exact form the apology takes depends on the severity of the offense, the social situation, and the social distance (low class, middle class, and higher class). The most important distinction is the severity of the offense:

1. The offense is minimal, such as touching a person inadvertently in a social setting (e.g., classroom or a meeting), or forgetting to fulfill a promise. In such situations the offender may say *aasif*, *ʿafwan*, *maʿthiratan*, *ʔtithaaraatii*, or *almaʿthirah*, "Please accept my apology" without any elaboration. The words *ʿafwan* and *maʿthiratan* are most likely to be used among nonacquaintances and strangers in formal linguistic formulas such as the following:

maʿthiratan yaa akhii	My apology, my brother!
maʿthiratan yaa ukhtii	My apology, my sister!

or

aasif yaa akhii	Sorry, my brother!
aasif yaa ukhtii	Sorry, my sister!

2. The offense is severe, such as hurting someone's feelings, embarrassing someone in front of a group of people, accusing someone falsely, or damaging property. In such social situations it is not enough to say, *aasif* or *ʿafwan*. The offender must add an elaboration and a facial expression connoting physical pain to show that the apology is sincere. Examples of linguistic formulas that can be used in such situations between friends and strangers alike are the following:

anaa muta?sif jidan [+ a sincere facial expression + detailed explanation]
(I am very sorry)

anaa muta?sif jidan wa?rjuu an tusaamiḥunii [+ a sincere facial expression + detailed explanation]
(I am very sorry and I beg your forgiveness)

min faḍlak iqbal/ yaa akhii iqbal ʔtithaarii [+ a sincere facial expression + detailed explanation]
(Please accept/ please, my brother, accept my apology)

As the previous examples make clear, it is not enough to use only an expression of apology or an expression of responsibility; rather both are required in situations where severe offenses occur. In the case of property damage, the offender must offer to repair the damage or compensate for it.

Requests

Using such variables as rank, age, social status, and social distance, Ervin-Tripp (1976) classified the English directives into six categories: (a) needs statements, (b) imperatives, (c) embedded imperatives, (d) permission directives, (e) nonexplicit question directives, and (f) hints. By analyzing the forms of request in Arabic, one can find examples from each category identified above. However, in Arabic society, the speaking rules that cover people's social behavior are different from those used in an English-speaking community. The following are examples of requests that native Arabic speakers use in their daily lives:

Imperatives

This form of request is used among family members and close friends, for example:

ʔiṭinii khamsuuna duulaaran. (a woman to her husband)
(Give me fifty dollars.)
Taʕaala allayla washar ʿindanaa (a person to his close friend)
(Come tonight and entertain with us.)

Embedded imperatives

Sometimes the phrase, *min fadlak*, "if you please", *alrrajaa*? or *arjuu*, "I beg you" is used with imperative forms directed to a person equal or lower in rank to make the request more polite. An example of this would be a dean at a university asking a division chair to reply to an applicant:

min faḍlak rud ʿalaa alductor ...waʔkhbirhu an yabʿatha lanaa muahilaatu (high to low in rank)
(please respond to Doctor ... and tell him to send us his credentials)

arjuu min ḥadratikum almusaʕadah fii haathaa almawthuuʕ (equal in rank)
(I beg your honor for help in this matter)

Thus these forms of address indicate that plurals may be used to show respect. For instance, an individual from whom something is requested will be addressed as a group (*min ḥadratikum* instead of *min ḥadratik*a, "he" or *min ḥadratiki*, "she" particularly in writing). The reason behind this usage is *min ḥadratik* seems so dry and less respectful than *min ḥadratikum.* In the following example in which a person addressing one of his acquaintances who witnessed someone beating his son, but did nothing to stop it:

uriidu min ḥadratika an laa tatadakhkhal, laqad kunta mawjuud hunaak ʿindamaa kaan yaḍribuhu.
(I want from you not to interfere, you were there while he was beating him.)

Permission requests

This type of request is used for asking permission to do something. The following are some examples of permission requests and who uses them with whom.

atasmaḥu lii yaa ustaath an athhaba waaghsil yaday (pupil to his teacher in the class)
(Could I have your permission to go and wash my hands?)

atasmahu lanaa yaa ustaath bi ahmad khamso daqaaiq.
(teacher to another teacher)
(Could we have your permission to take Ahmed for five minutes?)

tismah lii yaa abii an athhab maʿ aṣḥabii ilaa naablus limushaahadat almubaaraah (son to father)
(Could I have your permission to go to Nablus to watch the match?)

Other requests for permission, such as

mumkin astaʿmil alqalam
(May I use the pen?)

or

mumkin ilqalam daqiiqah
(May I have the pen for a minute?)

are used especially with nonacquaintances.

Nonexplicit question requests

Such requests are used among equals and non-equals in rank.

hal ʿindaka maanʿ ithaa ghayrnaa almawʿid ilaa yawm alithnayn alsaʿah altaasʿa ṣabaahan? (student to instructor, non-equals in rank)
(Would you mind if we change the date to Monday at 9:00am?)

hal tazuuranii ghadan? (friend to friend, equals in rank)
(Will you visit me tomorrow?)

866

Using *hal* at the beginning of a request as in the expression *hal ʿinddaka maanī* occurs in spoken Modern Standard Arabic and in writing only. In Jordanian colloquial Arabic the expression *ʿindak manī* is used instead.

Such requests are not used frequently by native Arabic speakers in comparison to native English speakers, who use this strategy extensively in their written and spoken discourse. Moreover, native English speakers sometimes use nonexplicit question requests as complaints, which are misunderstood by native Arabic speakers as simple requests for information rather than as complaints plus requests, for example:

S: Is this sandwich salted?
H: I don't think so.

Here the speaker (native American English speaker) is complaining to her husband (nonnative American English speaker) that he forgot to salt the sandwich. The husband, however, understands her question as a request for information rather than as a complaint that he should have salted the sandwiches plus a request to pass or bring the salt. Another significant strategy used by native speakers of English is the question format Can you close the door? or Could you close the door? as a polite request to close the door rather than a question of ability, as the surface meaning implies. This strategy is sometimes used in Arabic as well. An equivalent request would be *mumkin itsakir ilbaab* or *bitiqdar itsakir ilbaab*. The question *mumkin itsakir ilbaab?* does not in any way, on the surface or the deep level, imply a question of ability, whereas the question *bitiqdar itsakar ilbaab?* on the surface implies a question of ability, but on the deep level does not.

Hints

This type of request is used among close friends and family members, and sometimes with nonacquaintances, particularly, those whose presence in a situation is not welcomed or desired. For instance:

fii alwaqt alḥaadir ʿaanii min ḍaaiqah maaliyah.
(At the present time, I am having financial problems.)

This statement made in the presence of a friend or a family member in the course of a discussion of financial issues, implies the speaker's need to borrow some money.

Other hints are aimed at ridding oneself of guests. A host may yawn and say, *anaa taʿbaan wamurhaq jidan* (I am tired and very exhausted) hinting for the person(s) to leave. Another tactic is to offer the coffee early in the visit. In Arabic culture the coffee is always served at the end of the visit; thus offering the coffee early constitutes a hint to leave. Hosts may hint that they would like their guests to leave by not being actively involved in the conversation, acting oddly with family members, and making tired or sleepy facial expressions asking guests to leave. Examples of hints used by guests as a sign that they are ready to leave are yawning and stretching, asking for the

coffee, putting on the shoes, the jacket or the sweater, leaning toward the door and leaning forward. Breaking eye contact, which American native speakers use as a sign of leave-taking (Knapp et al, 1973), does not exist in the nonverbal behavior of Arabic native speakers. Americans learning Arabic and teachers of Arabic must at least be aware of this nonverbal behavior.

The Expression of Gratitude

Expression of gratitude in Arabic culture is a necessary sign of politeness and appreciation. A number of expressions are available, but each one is used in a different social situation. The following are some expressions of gratitude that native Arabic speakers use in their daily life interactions:

1. *shukran* (Thank you): Used formally and informally in all situations.
2. *shukran jaziilan* (Thank you so much): Used formally and informally in situations where the gift, favor or service requires a special effort.
3. *mercii* (Thank you): Borrowed from French used only by women in formal and informal situations.
4. *baaraka allaahu fiika* (May Allah [God] Bless you): Used formally and informally instead of the above three expressions, especially by religious people.
5. *jazaaka allaahu khayran* (May Allah [God] reward you): Used formally and informally in all situations, particularly by religious people.
6. *Allaah ikhalii iwlaadak* (May Allah protect and retain your children): Especially used by women in informal and formal situations.
7. *inshaa Allaah bilafraah* (By Allah's permission, always on happy occasions)
inshaa Allaah bitakharujak (By Allah's permission, on your graduation)
inshaa Allaah waallah mi'dikum alṣabii (By Allah's permission, when Allah grants you a son)
inshaa Allah bi'irsak or inshaa Allah bifarhitii fiik
(By Allah's permission, on the occasion of your wedding)
inshaa Allah wamaḥalukum ᶜaamir (By Allah's permission, while your house is full of happy occasions)
8. *akala ṭa'aamukum alabraar, waafṭara ᶜindakum alṣaaimuun, waṣalat ᶜalaykum almalaaikah, wathakarukum Allaahu fiiman ᶜindah* (May only the pious and fasting people eat your food, may the angels pray for you, and may Allah mention you amongst those who are in his presence). This is used only after eating at somebody's house. It is always used by religious people.

868

Although the above list is not an exhaustive one, it contains some of the most common forms of expressing gratitude. Students of Arabic as a foreign language are not required to learn all these expressions. However, they should learn Nos. 1, 2, and 5 because many people use them.

It is worth mentioning here that native Arabic speakers, often respond to expressions of gratitude by saying *haathaa min waajibna* (It's our duty). Such a response is used to express good manners, respect, and closeness of the relationship between the participants.

Greetings

In Arabic culture, the greeting is an essential element of people's daily lives. They are commanded to greet others and to answer a greeting with a better one. "When a courteous greeting is offered you, meet it with a greeting still more courteous or at least of equal courtesy." (*Holy Quran*, Chapter IV, Verse 86). Muslims are encouraged strongly to greet each other. A person is required to greet other people as long as they can hear him/her within a reasonable distance. No greeting is required if other people are visible, but out of earshot. However, a person might wave the right hand to greet people, especially acquaintances, too distant to hear a verbal greeting. The passing, walking person offers the greeting to those who are sitting. The following story illustrates this point interestingly.

> But the story is told of a British agent in Crete during World War II. He spoke fluent Greek and dressed like a Greek who was once sitting beside a road as an old woman approached. Cherete 'hello' he called out, but she walked on toward him and said in a low voice "God bless you, my son, and take you safe back to your native land." At length he figured out how she had penetrated his disguise: he had violated an unspoken rule that the moving person utters the first greeting (Applegate, 1975, p. 278).

Also, in Arabic culture seated people must stand up if the greeter has stopped to shake their hands. It is impolite and disrespectful to shake people's hands while sitting.

The most frequent Islamic greeting appropriate anytime and anywhere is *alsalaamu ʿalaykum* "Peace be with you." The response is *waʿalaykum alsalaam* or *waʿalaykum alsalaamu warahmatu Allahi wabarakaatuh,* "And peace be with you also," or "And peace be with you also and Allah's mercy and blessings." This greeting can be used anytime during the day, regardless of the relationship between the interlocutors. It can be also used to initiate a conversation.

In addition to the "Peace be with you" greeting, there are other expressions that learners of Arabic can use at different times during the day. These are (a) *ṣabaahu al-khayr,* "Good morning," (b) *masaa?u al-khayr,* "Good evening," and (c) *marhaba,* "Hi." *Marhaba* is used anytime, with acquaintances and nonacquaintances. It is found

more in the sociolinguistic behavior of people who are not very religious. It does not exist in Islamic culture and some scholars argue it is not appropriate at all. (d) *Allah yʿtiik alʿaafiyah,* "May Allah protect your health" is another greeting expression that is widely used by Arabic speakers. This linguistic formula is used only when someone is passing by a another person who is working on something. The correct reply is *Allaah yʿaafiik,* "May Allah protect your health, too." It is noteworthy here that some people use this greeting as an expression of gratitude, particularly for something that required physical and/or mental efforts.

Refusals

The speech act of refusal is a very sensitive issue for Arabs. In Arabic culture a person is strongly encouraged to comply with a request for help; to accept an invitation to a party or dinner or an offer of food, drink, or a gift; and to provide requested suggestions pertaining to a specific subject, if possible. If a person cannot comply, the appropriate linguistic refusal formula depends on the relative status of the interlocutors and on their social relationship. Analyses of the strategies for refusing used by native Arabic speakers revealed two categories: direct and indirect. Examples of direct refusals, which are used with intimates and status equals only are the following:

laa (no)
laa astatiʿ (I cannot.)
mish faaḍi (I am occupied.)

The interesting point here is that such refusing strategies are not used with strangers. In contrast. native speakers of American English have been found (Beebe et al, 1985) to say "no" or "I refuse" not only to intimates and status equals but also to strangers. If Arabs are asked something by a stranger they will help if they can, send him/her to somebody who can help, or they will say *anaa aasif jidan laa aʿrif,* "I am very sorry, I do not know." Native speakers learning Arabic must be aware of this fact; otherwise their linguistic behavior will be interpreted as rude and uncharitable.

The other category of refusing strategies, indirect refusals are used with acquaintances of equal status and with close friends of unequal status. Note the following strategy: an expression that indicates a strong desire to comply, inability to comply and an explanation.

atamanaa law astaṭiʿ, ʿindii moḥaaḍarah fii thaalika almasaa? (I wish I could, [but] I have a class that evening).

yaa rayt, anaa mashghuul jidan, inshaa Allaah fii almara alqaadimah (I wish so [i.e., to do it], [but] I am very busy, by Allah's permission next time.)

Similarly native speakers of American English were found using in sequence an expression of positive opinion, an expression of regret or an excuse, and a reason or an explanation (Beebe et al, 1985). The only difference between American and Arab refusals is that Arab people use this strategy not only with acquaintances of equal status but also with acquaintances of unequal status. According to Beebe et al., (1985) other indirect strategies used by Americans that are also utilized by Arabic speakers, include: an alternative statement (*atamanaa law hakaytlii min qabl,* "If you had asked me before"), a statement of principle (*anaa laa ataʿaamal maʿ aṣḥaabii bilfuluus*, "I do not do business with friends"), an unspecified or indefinite reply (*inshaa Allaah sauʔḥaawil*, "By Allah's permission I will try"), verbal or nonverbal avoidance such as silence or a change of topic.

Partings

Strategies for ending a conversation or a meeting in Arabic culture vary with the status of the participants and the relationship between them. Whether the meeting is formal or informal no one may leave without letting the participants know. Leaving without taking leave is considered ill-mannered, disrespectful and indicative of a flawed personality. Regardless of status differences, if the participants have a close relationship, leave-taking strategies are informal. Note the following strategies:

1. *anaa laazim amshi, inshaa Allaah araakum ghadan*
   (I have to go, by Allah's permission I will see you tomorrow)

2. *tuṣbihuuna ʿalaa khayr*
   (I hope you wake up in the morning [only used at night]
   and everything is fine.)

3. *astawdiʿukum Allaahu alathii laa taḍiʿu wadaaʿuh*
   (I leave you with the trust of Allah, whose trust has never been
   betrayed)

4. *naraakum inshaa Allaah ghadan, alusbuuʿ alqaadim, . . .*
   (See you, Allah willing, tomorrow, next week, . . .)

Those staying behind respond by saying, *mac alsalaamah,* "Peace will be your companion." But the response to *tusbihuuna ʿalaa khayr*, is *waantum min ahlu*, "and you too." But the statement *maʿ alsalaamah* is always used, even if other responses are required, in which case it follows the required expression. An interesting point here is that some of these statements are marked. That is, they are used almost exclusively by practicing Muslims. An example of this is statement number three above.

In addition to the aforementioned verbal strategies, there are other nonverbal behaviors used as signs of asking people to leave, in particular. Such strategies are discussed on pages 76 and 77 under the subtitle *Hints*.

Disagreement

Strategies for disapproving an action or a deed in Arabic culture vary depending on the persons involved, the social distance between them, their status, their ages, and the subject of disagreement. Among intimates, direct disapproval is expressed in most cases, for example:

anaa ghayr muwaafiqh	(I do not agree.)
anaa laa atafiqu maʿaka fii haathaa	(I do not agree with you on this.)
anaa akhtalifu maʿaka fi haathaa	(I disagree with you on this.)

In addition to such declarative expressions, imperatives are also used to express disagreement, particularly with those the speaker has authority over, such as children, and a teacher with his/her students. An example of this would be a teacher saying to a student who was found writing on the board during a break: *imsaḥ allawḥ wathhab waintaṭirnii fii maktabii,* "Clean the board and go and wait for me in my office." Interrogatives are sometimes used to express disagreement. An example of this would be *maathaa taʿmal hunaa,* "What are you doing here?" as one person said to his close friend when he found him playing when he should have been studying for his final exam.

Disagreement with close friends and family members might be expressed using a proverb. For instance, note this remark made by a father to his son, who was laughing in the presence of a group of friends in the middle of a conversation: *iḍḍiḥik min duun sabab qilit adab*, "Laughter without (a) cause is a sign of impoliteness." Using such statements with anyone but intimates will create major difficulty between the interlocutors. The age of the person to whom such statement is addressed is very significant even if he/she is a son or daughter. In brief this proverb and other similar insulting ones are used with children in family settings or by teachers with their pupils in the classroom; only rarely, are they uttered in public to adults.

Disagreement with adults in family settings, in public discussions, and in formal business meetings sometimes takes nonverbal form. Such disapproval is expressed by staring at the speaker with a facial expression indicating disagreement or anger, or sometimes by leaving the room or the meeting while the person is speaking. The latter is used only among family members and in informal meetings with friends and acquaintances. Such behavior is not considered polite, but still takes place. A word of caution is needed here regarding looking at people while they are talking, because native English speakers interpret this as a sign of interest, approval, and respect for the speaker. In Arabic culture, such behavior occurs only when something is going wrong in the discussion. Native English speakers should keep this in mind while dealing with native Arabic speakers.

Telephone Etiquette

The organization of discourse in speech acts such as telephone calls is covered by social rules for a proper beginning, middle, and end. "Violations of such rules . . . [which] vary so much from one culture to another . . . and misperceptions on the part of nonnative speakers who operate under different cultural assumptions and attitudes are likely to result in severe cross-cultural misunderstandings" (Schmidt, 1988, p. 64). To shed some light on the discourse sequence of Arabic telephone calls, the following example is provided:

Caller:	(Dials and the telephone rings.)
Answerer:	*alloo* (Hello)
Caller:	*alsalaamu ʿalaykum* (Peace be with you)
Answerer:	*waʿalaykum alsalaam, ahlayn, min? Anwar, ahlan yaa Anwar keif haalak, inshaa Allah mabsuuṭ*, (Peace be with you too, who? Anwar, you are welcome Anwar How are you, I hope you're happy, by Allah's will)
Caller:	*alhamdulilaah* (Thank God)
Answerer:	*kayf ṣihatak* (How is your health?)
Caller:	*alhamdulilaah* (Thank God)
Answerer:	*kayf haal zawjatak wabintak, inshaa Allah bikhayr*. (How are your wife and daughter? I hope they are well by God's will.)
Caller:	*alhamdulilaah* (Thank God). *kayf haalkum intum, insha Allaah mabsuḍiin.* (How are you, by God's will I hope your are all well.)
Answerer:	*alhamdulilaah, kuluna bikhayr* (Thank God, all of us are well.)
Caller:	*ʿami abuu bashaar mawjuud?* (Is my uncle Abu Bashar there?)
Answerer:	*yaa, lahthah ishway* (Yes, just a moment.)
Answerer:	*allo* (Hello)
Caller:	*kayf alhaal yaa ʿami, inshaa Allaah mabsuuṭ waṣihatak bikhayr.* (How are you? I hope you are fine and your health is good by God's will.)
Answerer:	*alhamdulilaah* (Thank God). (the talk goes for 1-3 minutes)
Caller:	Starts the subject for which he/she called, the talk continues until they finish the topic.

After discussing the topic, the interlocutors start using what Marcus and Slansky (1994) call "softening" statements indicating their intention of ending the telephone conversation. The call ends in most cases by using the expression "*maʿ alsalaamah.*"

By comparing this telephone call with an American one, we can detect the following differences that cause misunderstanding and irritation. First, American callers identify themselves to the answerer (e.g., "This is professor Robert Fenske calling from ASU . . ."). In contrast, in Arabic telephone calls neither the caller nor the answerer identifies him/herself. Rather, each must try to guess the other's identity. Only

if they fail, do they identify themselves. Second, sometimes Americans ask to speak with someone other than the answerer even though they know the answerer and might be a close friend or relative (e.g., "May I speak to Cindy, please!"). To a Middle Easterner, such behavior is considered not only strange but also offensive, rude, and disrespectful. Moreover, the answerer interprets this to mean that there is something wrong between him/her and the caller that he/she is not aware of.

Moreover, Americans dealing with Arab people should be aware that women, especially young unmarried ones, refuse to identify themselves if they have been asked to by male callers. This occurs in particular when the caller is a stranger who has identified himself. Usually women identify themselves by saying *anaa ibnatuhu*, *anaa ukhtuhu*, or *anaa zawjatuhu,* "I am his daughter," "I am his sister,"or "I am his wife," respectively. Under no circumstances will a young female tell a caller her name. But married women will identify themselves as, for instance, *um Ahmed* or *zawjat ʿamar* ("The mother of Ahmed," or "Ammar's wife" respectively). Admittedly, such behavior seems strange and is discomforting to American males calling Arab residences. However, it is simply one of the cultural differences Americans must be aware of in dealing with others.

Pedagogical Implications

Based on the foregoing discussion of the speech acts in Arabic language, a number of pedagogical implications can be inferred for language learners, language instructors, and curricula developers:

1. Textbook writers must include cross-cultural content in their books. In cases where such information is not included in the text, teachers can find information about the unwritten rules of speech acts in books designed specifically for international trade and marketing communication. Examples of such books are Dulekk and Fielden (1990) and Murphy and Hilderbrandt (1991).
2. Experienced teachers who have considerable experience in the native language culture and the target language culture must be hired wherever possible. Such teachers are capable of personalizing the material and adding to it based on their firsthand experience.
3. Mass media is an interesting and wonderful instrument that can provide students with a variety of linguistic and nonlinguistic features. For example, television programs and films are a good source of input for the unwritten rules of speech acts. Moreover, teachers and students can videotape interviews with people from the target culture and present them in the classroom for observing and scrutinizing the linguistic and nonlinguistic behavior of the native speakers of the target language.

4. Students must be encouraged to think about and compare their linguistic behavior with the linguistic behavior of the target language group. Discussion of the similarities and differences between the two cultures will be very fruitful and interesting.
5. Language instructors must not teach the unwritten rules of speech in an "ethnocentric manner." In other words, they should be very careful not to suggest that any particular culture's rules of speech are right or wrong. Increasing awareness of similarities and differences for optimal communication should be the goal of language learning and training.
6. Language learners and language teachers alike must be aware of stereotypes based on a few observations of target language sociolinguistic behavior. For instance, some might say Arabs are "too physical;" a more objective and accurate statement would be that close friendship between Arabs requires less space and more touching than friendships between Americans. A stereotype that Arabs hold about Americans is that they are not generous in offering a lot of food when inviting an Arab to dinner; a more objective statement would be that they are economical and try not to waste food by cooking more than the diners are likely to want.
7. Pointing out and discussing those speech acts that are likely to annoy and baffle Arabic speakers on the one hand and those that annoy and baffle English speakers on the other—is a good way to make students aware and tolerant of the other people's way of speaking and interacting. Examples of what annoys and baffles Arabs about Americans are: the lack of religious expressions (such as by "God's will," "God forbid") in their speech; saying "oh" or "wow" when they see something exceptionally beautiful or good; using the interlocutor as an example when speaking about undesired things (e.g., an American said to an Arab while passing by a cemetery, "That is where they put you when you are dead"); using intimate expressions such "Honey," or "My love" in the presence of other people; and saying something that is completely different from what is in their minds or hearts. Examples of what annoys and baffles Americans about Arabs are the constant reference to God in their speech (e.g., "Thank God," "by God's will"); asking inappropriate personal questions such as "How old are you?" or "How much money do you make?"; and making comments on people or things (e.g., "Why are you very fat? Do not eat," advised an Arab his American female friend.).
8. "Teachers should develop a student's metapragmatic ability" (Thomas, 1983, p. 98). That is, they should equip the student with the skills necessary to analyze language use consciously. Such metapragmatic ability might be achieved by discussing language use in the light of the parameters proposed by Van Dijk (1977, p221): (a) positions (e.g., roles or status), (b) properties (e.g., sex or age), (c) relastions, (d) functions (e.g., father, waitress, or judge).

9. "[It is] advisable for the teacher to explicitly point out to the learner that polite ness and other speech acts are an integral part of the foreign culture system, and should be neither used nor interpreted by reference to the learner's native system. More effective teaching of the behavioral component may minimize native cultural interference and prevent impolite, ineffective, or otherwise inappropriate behavior on the part of the learner" (House & Kasper, 1981, p.184).
10. Teachers in particular, but also learners, should be aware of the difference between pragmalinguistic failure and sociopragmatic failure. That is, they should know "the first is language specific and it should be possible for the teacher to correct it quite straightforwardly. The second is in part culture-specific, a reflection of the student's system of values and beliefs, and not be corrected but only pointed out and discussed" (Thomas, 1983, p.109).
11. Instructors should strongly encourage their students to have a direct and meaningful experience with the surrounding local community of the target language. Personal experiences with the people of the target culture can lead to a better understanding of the students' own culture through comparison and contrast with the target culture. Additionally, intercultural experiences help reduce the anxiety in unfamiliar cross-cultural encounters. The reduced anxiety boosts self-confidence, increases cultural competency, and improves cross-cultural communication.
12. Instructors and learners should be aware that "cultural competency entails personal and interpersonal skills. Among the personal skills are: (a) the capacity to become aware of one's own cultural perspective and interpersonal behavior, and (b) the ability to interpret cultural exchanges accurately. Interpersonal skills include withholding judgement until one knows others better and understands their culture and skillful interpretations of cross-cultural nonverbal communication" (Chisholm, 1994, p. 60).

References

Applegate, R. B. (1975). The language teacher and the rules of speaking. *TESOL Quarterly, 9*, 271-281.

Barlund, D. (1968). *Interpersonal Communication: Survey and studies*. Boston: Houghton Mifflin.

Barna, L. R. (1983). The stress factor in intercultural relations. In D. Landis & R. W. Brislin (Eds.), *Handbook of intercultural training* (Vol. 2). New York: Pergamon Press.

Beebe, L. M. et al. (1985). Pragmatic transfer in ESL refusals. In R. Scarcella, E. Anderson, & S. Krashen (Eds.), *On the development of communicative competence in a second language*. Rowley, MA: Newbury House.

Berry, J. W. (1970). A functional approach to the relationship between stereotypes of familiarity. *Australian Journal of Psychology, 22*, 121-128.

Bonivillain, N. (1993). *Language culture and communication: The meaning of messages*. Englewood Cliffs, NJ: Prentice Hall.

Brewer, M. B., & Campbell, D. T. (1976). *Ethnocentrism and intergroup attitudes: East African evidence.* New York: Wiley/Halted.

Brown, R. W., & M. Ford. (1961). Address in American English. *Journal of Abnormal and Social Psychology, 62*, 375-385.

Curt, C. J. (1976). *Nonverbal communication.* Bronx, NY: Northeast Center for Curriculum Development.

Davidson, A. R., & Thomson, E. (1980). Cross-cultural studies of attitudes and beliefs. In H. C. Triandis & R. W. Brislin (Eds.), *Handbook of cross-cultural psychology: Social psychology* (Vol. 5). Boston: Allan and Bacon.

Delgato-Gaitan, C., & Trueba, H. T. (1985). Ethnographic study of participant structures in task completion: Reinterpretation of "handicaps" in Mexican children. *Learning Disability Quarterly, 8,* 67-75.

Dulek, R. E., & Fielden, J. S. (1990). Principles of business communication. New York: Maxwell Macmillan.

Ervin-Trippe, S. (1976). Is Sybil there? The structure of American English directives. *Language in Society, 5*(1), 25-66.

Giles, H., Coupland, N., Williams, A., & Leets, L. (1991). Integrating theory in the study of languages. In R. L. Cooper & B. Spolsky (Eds.), *The influence of language on culture and thought: Essays in honor of Joshua A. Fishman's sixty-fifth birthday* (pp. 113-136). New York: Mouton de Gruyter.

Hall, E. T. (1973). The silent language. Garden City, NY: Anchor Books.

House, J., & Kasper, G. (1981). Politeness markers in English and German. In F. Coulmas (Ed.), *Conversational routine* (pp. 157-186). The Hague: Mouton.

Huitt, W. G. (1988). Personality differences between Navajo and non-Indian college students: Implications for instruction. *Equity and Excellence, 24*, 71-74.

Jacobs, R. L., Sr. (1990). Learning styles of black average and low achievers. *The Clearing House, 63*, 253-4.

Jacobs, R., & Jacobs, J. (1988). Individual and cultural differences in learning styles. *The Journal: The International Association of Pupil Personnel Workers, 32*, 45-148.

Janicki, K. (1985). On the feasibility of pedagogical contrastive sociolinguistics. In Jack Fisiak (Ed.), *Contrastive linguistics and the language teacher.* Oxford: Pergamon Press.

Knapp (1973). The rhetoric of good-bye: Verbal and nonverbal correlates of leave-taking. *Speech Monographs, 40*, 182-198.

LaFrance, M., & Mayo, C. (1978). *Moving bodies: Nonverbal communications in social relationships.* Boston: Houghton Mifflin.

Landis, D., Brislin, R., & Fischer, G. W. (1985). Attributional training versus contact in acculturative learning: A laboratory study. *Journal of Applied Social Psychology, 15*, 466-482.

Marcus, S., & Slansky, N. (1994). Teaching the unwritten rules of time and space. *ELT Journal, 48*, 306-314.

Murphy, H., & Hildebrandt, H. (1991). *Effective business communication.* Singapore: McGraw Hill International.

Schmidt, R. (1988). Applied sociolinguistics: The case of Arabic as a second language. *Anthropological Linguistics, 27,* 94-101.

Stephan, C. W. (1992). Intergroup anxiety and intergroup interaction. In J. Lynch, D. Modgil, & S. Modgil (Eds.), *Cultural diversity and the schools: Prejudice, polemic or progress?* London: Falmer Press.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues, 41,* 157-176.

Stephan, W. G., & Stephan, C. W. (1989a). Antecedents of intergroup anxiety in Asian-Americans and Hispanic-Americans. *International Journal of Intercultural Relations, 13,* 203-216.

Stephan, W. G., & Stephan, C. W. (1989b). Emotional reactions to interracial achievement outcomes. *Journal of Applied Social Psychology, 19,* 608-621.

Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics, 4,* 91-104.

Van Dijk, T. A. (1977a). Context and cognition; Knowledge frames and speech act comprehension. *Journal of Pragmatics, 1,* 211-232.

Wolfson, N. (1990). The social dynamics of native and nonnative variation in complimenting behavior. In M. Eisenstein, (Ed.), *Variation in second language acquisition*: Empirical views. New York: Plenum Press.

Author

ANWAR A. HUSSEIN, Assistant Professor, Professional Education Program, Ottawa University, Tempe Center, 4545 S. Wendler Dr. Tempe, AZ 85282-6308. Specializations: foreign/second language acquisition and grammar, foreign/second language pedagogy, sociolinguistics.

Reviews

Discourse (1989, 1992). By Guy Cook. Teacher Education Series. Christopher Candlin and Henry Widdowson, Eds. Oxford: Oxford University Press.

Reviewed by MARTHA S. BEAN, *San José State University*

Discourse (now in its third printing) reflects a growing trend on the part of teacher training programs to consider in language analysis oral exchanges and written texts that are longer than sounds, words, and sentences. Cook begins by defining discourse as natural language in use, pointing out that knowledge of discourse form and function is crucial to the communicative success of the nonnative speaker. Having established a rationale for teacher learner concern with discourse, Cook organizes the ensuing discussion around three sections: (1) a survey of theories of discourse, (2) a demonstration of the workings of discourse in language learning and teaching, and (3) an exploration of the ways in which teachers can use both theoretical and classroom-centered approaches. Within this framework, teachers first work with background information on discourse (a theoretical approach), then observe how discourse operates within the presentation, practice, and content of the language learning setting (a pedagogical approach), and finally selectively adjust their classroom practices according to their heightened awareness of discourse style and function (an action-oriented approach). Each section includes multiple tasks that highlight the issues presented in the text.

In section one (pp. 3-75), Cook first contrasts Chomsky's work with the work of Hymes and Halliday, reiterating the learner's need for the study of "language in use." After unpacking such cohesive devices as consistency in verb form, parallelism, etc., Cook proceeds to the pragmatic considerations of speech act analysis, reiterating that knowledge of linguistic form alone is insufficient for communicative competence. Cook next contrasts the work of Sinclair and Coulthard with the work of Sacks, Schegloff, and Jefferson—showing how the former have uncovered a ranked structure within lessons, while the latter have uncovered both linguistic and non-linguistic mechanisms of turn-taking. Finally, Cook treats discourse as dialogue by examining the information structure of selected segments and, in a helpful review of the literature on schemata (pp. 70-74), by discussing how language users cognitively organize incoming chunks of language to comprehend and retain them.

Sections two and three (pp. 79-155) provide teachers with useful exercises through which they can discern the underlying motives of language teaching texts while helping their students become more effective receivers and senders of discourse in the target language. In section two, Cook initially contrasts bottom-up with top-down approaches to handling text and emphasizes the importance for learners of knowing as much as possible about their audience. He next deals with the "parts" and functions of

spoken and written discourse types while inviting teachers to reflect on the interplay of discourse type and topic in tasks based on segments from ESL/EFL textbooks. Finally, he addresses the teaching of conversational skills from a discourse analytic perspective, demonstrating how teachers can do consciousness-raising work with their learners on linguistic devices operating at the discourse level. Section three presents tasks for teachers and learners that apply the information of sections one and two. These tasks effectively engage teachers and learners in reflection on and application of the discourse concepts addressed earlier in the text.

Cook's work, presented largely in a conversational tone, is very accessible to teachers and provides an excellent overview of the field of discourse analysis without bogging down in excessive jargon and particular perspectives. Teachers hungry for a user-friendly theoretical overview of discourse analysis will relish section one, while those interested in applying a discourse analytic approach to language-teaching contexts will find sections two and three invaluable. The latter, useful as they are, could have been even further strengthened by including more tasks aimed directly at learners, such as the engaging train-story riddle at the end of section two. Readers will find their use of the text additionally facilitated by a detailed table of contents, a glossary of terms, and an extensive index.

Only one major area of omission emerges. Teachers will find little information on the discourse-styles of non-western cultures (beyond a cursory mention on p. 98) and on the ways in which teachers and learners might work together to accommodate each other's discourse styles, should they deem it desirable to do so—the kinds of insights yielded by the work of Scollon and Scollon (1981), and Tharp and Gallimore (1988). Thus, while Cook's book helps teachers to understand discourse and discourse analysis better in terms of teaching tasks and texts, it only implicitly addresses the issue of teacher-learner mismatch in discourse style (Kumaravadivelu, 1991) and gives short shrift to tasks in which learners might investigate the discourse styles of their classrooms and social circles as well as of their texts and exercises. In sum. What the book does, it does well; it simply stops one section short.

References

Kumaravadivelu, B. (1991). Language learning tasks: Teacher intentions and learner interpretations. *English Language Teaching Journal, 45*(2), 98-107.

Philips, S. U. (1983). *The invisible culture: Communication in the classroom on the Warm Springs Indian Reservation.* New York: Longman.

Scollon, R. & Scollon, S. B. K. (1981). *Narrative, literacy and face in interethnic communication.* Norwood, NJ: Ablex.

Tharp, R. & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning and schooling in social context.* Cambridge: Cambridge University Press.

Introducing Language Awareness (1995). By Leo Van Lier. Penguin English Applied Linguistics Series. Ronald Carter and David Nunan, Eds. London: Penguin. Pp. 161.

Reviewed by LIDIA WOYTAK, *Defense Language Institute*

This textbook is a part of a series of guides to four basic language skills—grammar, discourse analysis, standard English, and classroom interaction. These guides aim to help teachers and trainees keep abreast of applied linguistics and its relevance to common problems of teaching.

Leo Van Lier writes that the aim of his guide is to raise awareness of language use which he defines as "a person's sensitivity and conscious awareness of language and its role in human life" (p. xi). The term *language awareness* is relatively new. It was created in Great Britain in response to dissatisfaction with teaching by a group of educators who called for "awareness of language" or "knowledge about language" in the curriculum. Editors of a journal *Language Awareness* elaborate further: "it rose mistily from the consciousness-raising movements of the seventies, and was tidied up and documented in the eighties. The nineties should broaden it out and fatten it up" (p. 1). It is noteworthy that Leo Van Lier's understanding of language awareness goes beyond the scope of a classroom. It also "includes an awareness of power and control through language and of the intricate relationship between language and culture" (p. xi).

Van Lier designed *Introducing Language Awareness* as a reference and a workbook. Throughout this textbook the author acquaints readers with key works in applied linguistics and designs for them a variety of activities. In many instances the author addresses the reader directly. Thus, an activity on building blocks of language he starts with, "I invite you to list all the units or 'pieces' of language you can think of..." He skillfully draws examples from such mundane occurrences as car repair or a bike ride.

In "*Language Awareness: What, When and Why,*" Van Lier differentiates between focal and subsidiary awareness. He alerts readers to situations in which language is an intricate part of a problem. According to the author, such problems can be solved by focusing directly on the language. This approach is reminiscent of ideas expressed by S. I. Hayakawa in *Language in Thought and Action.*

In the chapter on "Building Blocks of Language" Van Lier attempts to identify elements of "recipe" knowledge for a language user. To him language, "is like a window which we only notice when it gets dirty. However, unlike a window which may be kept spotlessly clean just with a rag and some cleaning liquid, language seems to have inherent blemishes and weak spots which we continually have to work to overcome and patch up" (p. 12).

In the next chapter, "Context and Interpretation, " Van Lier discusses linguistic, interactional, and social contexts of language use. He guides the readers through the layers of discourse and then comparing it to using an onion, he illustrates

that meaning cannot be identified by focusing on a single layer. He concludes that the meaning of language is to a large extent determined by its use in actual context. The dependence of language on context was originally discovered by Bronislaw Malinowski, the anthropologist, who realized that despite the fact that he knew vocabulary and grammar of the Trobriand Islanders, he still could not communicate with them. Thus Malinowski concluded that to converse with the islanders, he needed to get involved in their experience. In volume two of *Coral Gardens and Their Magic* (1935, p. 11), he wrote "not even a sentence can be regarded as a full linguistic datum. To us, the real linguistic fact is the full utterance within its context of situation."

In the chapter on "Simmering Pots of Meaning" Van Lier introduces such concepts as connotation, collocation and predictability, metaphor, and transparency. He points out that many concepts can be expressed on a continuum from nice, through normal, to rude. In the same chapter he presents the table of continua of convention and creativity which divide the field into four textual areas. For each area he identifies several types of communication. Such a table could be very useful to a curriculum developer or a language teacher.

In the chapter on "Correctness" Van Lier discusses grammatical tolerance, differences between spoken and written language, and dialects. The author points out to the shallowness of the correct versus incorrect mindset and describes his "grammar stick" which indicates "shades of acceptability in context" (p. 83). Van Lier writes that spoken language, especially conversation, deserves more attention in foreign language instruction. The author reminds teachers that students need frequent and consistent practice in spoken language.

Van Lier states that language awareness plays a crucial role in educational settings. He writes, "If we want to educate, we need to know our language inside out, and play it with virtuosity and authenticity" (p. 100). Following a review of activities analyzing bits of language in isolation, he encourages educators to develop innovative language-awareness projects. Van Lier encourages them to focus on a crucial element of language in education—classroom interaction. Subsequently he presents a few samples of classroom talk which make readers aware of the use of target and native languages in the classroom. Next Van Lier complains that zealous educators in their race to conform to a multitude of standards, produce boring and unwieldy textbooks.

In the final chapter, "Critical Language Issues," Van Lier first reviews the contemporary literature and then proceeds to discuss the relationship between language and power. Through a variety of authentic examples, Van Lier demonstrates that frequently people use language to manipulate actions and thoughts of others. He guides readers to detect samples containing such language. This reviewer, for example, found an article on county staff reports (*Herald*, 15 June 95) which demonstrates a tense relationship between language and power.

> After the meeting, Johnsen said she is increasingly irritated by staff reports using "language that is obfuscating rather than clarifying." She suggested that some departments "may have the assumption that we're just going to approve the recommendation, sort of rubber-stamp it, no matter what it says, and I object to that." (p. 4C).

Van Lier's attitude towards bilingualism is positive. He states that "...every child has the right to speak, maintain and develop his or her mother tongue... Further he believes that..."an equitable educational sector has the duty to support this right actively" (p. 128). In *Language Awareness in the Classroom* (1991, p. 11), Carl James and Peter Garrett broaden up the issue of bilingual education by pointing out that, "there is some danger either of patronizing the children of some minority cultures or of exploiting them to the advantage of their native language (NL) English mainstream culture peers." Since bilingualism is a mushrooming issue especially in California schools where minorities of all backgrounds are becoming a majority, *multilingual* instead of *bilingual* would be a more equitable, accurate term.

In summary, readers of *Applied Language Learning* will find *Introducing Language Awareness* a welcome addition. Step by step, it reveals crucial elements of language by addressing directly concerns of foreign language teachers. Finally, the book reflects the views expressed in it: it is succinct, open, and convincing.

General Information

Calendar of Events*

1996

23-26 March, *American Association for Applied Linguistics*, Chicago. Information Elinor Ochs, Department of TESL and Applied Linguistics, UCLA Box 951531, 3300 Rolfe Hall, Los Angeles, CA 90095-1531; (310)825-0984, Email [ochs@humnet.ucla.edu].

26-30 March, *Teachers of English to Speakers of Other Languages*, Chicago. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703)836-0774, Email [conv@tesol.edu].

28-31 March, *Central States Conference on the Teaching of Foreign Languages,* Louisville. Information CSCTFL, Madison Area Technical College, 3550 Anderson Ave., Madison, WI 53704; (608)246-6573, Fax (608)246-6880.

11-13 April, *Southeastern Council on Latin American Studies,* Miami. Information Melvin S. Arrington, Jr., Modern Languages, University of Mississippi, University, MS 38677; (601)232-7713, Fax (601)232-7033.

11-13 April, *Southwest Conference on Language Teaching,* Albuquerque. Information Joann Pompa, Mountain Pointe High School, 4201 E. Knox Road, Phoenix, AZ 85044; (602)759-8449 Ext. 3036, Fax (602)759-8743.

11-13 April, *Languages and Communication for World Business and the Professions,* Ypsilanti. Information EMU Conference on Languages for World Business, World College, 307 Goodison Hall, Eastern Michigan University, Ypsilanti, MI 48197; (313)487-2414, Fax (313)485-1980.

11-14 April, *Association for Asian Studies,* Honolulu. Information Association for Asian Studies, 1 Lane Hall, University of Michigan, Ann Arbor, MI 48109.

18-20 April, *Russian Modernism and its European Context,* Lubbock. Information Lewis Tracy, Texas Tech University, Department of Classical and Modern Languages and Literatures, Box 42071, Lubbock, TX 79409-2071; (806)742-3286, Email [pvlew@ttacs1.ttu.edu].

18-21 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

22-24 April, *Language Classrooms of Tomorrow: Issues and Responses,* Singapore. Information The Director (Attention: Seminar Secretariat) SEAMEO Regional Language Centre, 30 Orange Grove Road, Singapore 1025, Republic of Singapore; (65)737-9044, Fax (65)734-2753, Cable: RELCENTRE Singapore, Email [tkhng@technet.sg].

25-27 April, *Pacific Northwest Conference on Foreign Languages,* Tacoma. Information Ray Verzasconi, PNCFL, Foreign Languages and Literatures, Oregon State

* Courtesy of *The Modern Language Journal* (University of Wisconsin Press)

University, Corvallis, OR 97331-4603; Email [verzascor@cla.orst.edu].

26-27 April, *Classical Association of the Atlantic States,* Baltimore. Information Sister Therese Marie Dougherty, Notre Dame College of Baltimore, 4701 N. Charles Street, Baltimore, MD 21210; (310)435-0100.

26-28 April, *California Foreign Language Teachers Association,* Sacramento. Information Sylvia Jones, 62 North Arroyo Boulevard, Pasadena, CA 91105.

9-11 May, *Cincinnati conference on Romance Languages and Literatures,* Cincinnati. Information Department of Romance Languages and Literatures, University of Cincinnati, Mail Location 0377, Cincinnati, OH 45221-0377; (513)556-2577.

27 May-1 June, *13th Annual Symposium, "Distance Learning,"* Albuquerque. Information CALICO, Duke University, 014 Language Center, Box 90267, Durham, NC 27708-0267; (919)660-3180, Fax (919)660-3138, Email [calico@acpub.duke.edu].

22-24 June, *Language Rights,* Hong Kong. Information Phil Benson, Department of English, Hong Kong Polytechnic University, Hung Hom, Hong Kong; Fax (852)2333-6569, Email [egphil@polyu.edu.hk].

24 June-26 July (dates tentative), *New Technologies & Less Commonly Taught Languages,* Honolulu. Information 1996 NFLRC Summer Institute, University of Hawaii at Manoa, East-West Road Building 1, Room 6A, Honolulu, HI 96822; (808) 956-9424, Fax (808) 956-5983, Email [nflrc@hawaii.edu].

July dates to be announced, *American Association of Teachers of French,* France. Information AATF, 57 E. Armory Avenue, Champaign, IL 61820.

8-13 July, *Esperanto Summer Workshop,* Hartford. Information Director of Summer Programs, University of Hartford, 200 Bloomfield Ave., West Hartford, CT 06117-1599; (800)234-4412.

18-21 July, *Society for the History of Authorship, Reading and Publishing,* Worcester. Information SHARP 1996, American Antiquarian Society, 185 Salisbury Street, Worcester, MA 01609-1634; Fax (508)754-9069, Email [cfs@mark.mwa.org].

4-9 August, *International Association of Applied Linguistics,* Jyvskyl, Finland. Information Udo O.H. Jung, Language Centre, University of Bayreuth, D-95440 Bayreuth, Germany; (0921)553595, Fax (0921)553594, Email [udo.jung@uni-bayreuth.de].

29-31 August, *EUROCALL 96,* Szombathely, Hungary. Information June Thompson, CTI Centre for Modern Languages, University of Hull, Hull HU67RX, United Kingdom; Email [eurocall@hull.ac.uk].

17-18 October, *Minnesota Council on the Teaching of Foreign Languages,* location to be announced. Information Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

18-20 October, *New York State Association of Foreign Language Teachers,* Kiamesha Lake. Information Helene Combopiano, 71Grand Boulevard, Binghamton, NY 13905.

18-20 October, *Texas Foreign Language Association,* Houston. Information TFLA, 1320 Modiste Drive, Houston, TX 77055; (713)468-4959.

23-25 October, *Foreign Language Association of North Dakota,* Bismarck. Information Herbert Boswau, P.O. Box 8198, Grand Forks, ND 58202-8198.

24-26 October, *Foreign Language Association of North Carolina,* Greensboro. Information Wayne Figart, PO Box 739, Wilmington, NC 28402-0739; (910)763-5733.

24-27 October, *Second Language Acquisition and Teaching,* Tucson. Information SLAT, Modern Languages 347, University of Arizona, Tucson, AZ 85721; Email [SLRF96@ccit.arizona.edu].

1-2 November, *National Association of Self-Instructional Language Programs,* Washington. Information NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19122.

22-24 November, *American Council on the Teaching of Foreign Languages with American Association of Teachers of German,* Philadelphia. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6810; (914)963-8830, Fax (914)963-1275.

27-30 December, *Modern Language Association of America,* Washington, D.C. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212)477-9863.

1997

6-8 March, *Southern Conference on Language Teaching with South Carolina Foreign Language,* Myrtle Beach. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912)333-7358, Fax (912)333-7389, Email [lbradley@grits.valdosta.peachnet.edu].

11-15 March, *Teachers of English to Speakers of Other Languages,* Orlando. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703)836-0774, Fax (703)836-7864, Email [tesol@tesol.edu].

3-6 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

10-13 April, *Central States Conference on the Teaching of Foreign Languages,* Columbus. Information CSCTFL, Madison Area Technical College, 3550 Anderson Avenue, Madison, WI 53704; (608)246-6573, Fax (608)246-6880.

5-7 August, *Problems of Teaching Modern Hebrew,* Jerusalem. Information Ben-Zion Fischler, Council on the Teaching of Hebrew, P.O.B. 7413, Jerusalem 91073, Israel.

7-11 October, *Foreign Language Association of North Carolina,* Winston-Salem. Information Wayne Figart, P.O. Box 739, Wilmington, NC 28402-0739; (910)763-5733.

21-23 November, *American Council on the Teaching of Foreign Languages,* Nashville. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914)963-8830, Fax (914)963-1275.

27-30 December, *Modern Language Association of America,* Toronto. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212)477-9863.

1998

17-21 March, *Teachers of English to Speakers of Other Languages,* Seattle. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2705; (703)518-2521, Fax (703)836-7864, Email [jennifer@tesol.edu].

September dates to be announced, *International Congress of Slavists,* Cracow. Information Robert Rothstein, Department of Slavic Languages, Herter Hall, University of Massachusetts, Amherst, MA 01003-3940.

27-31 October, *Foreign Language Association of North Carolina,* Winston-Salem. Information Wayne Figart, P.O. Box 739, Wilmington, NC 28402-0739; (910)763-5733.

20-22 November, *American Council on the Teaching of Foreign Languages,* Chicago. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914)963-8830; Fax, (914)963-1275.

27-30 December, *Modern Language Association of America,* location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212)477-9863.

The 20th Annual *Conference on the Teaching of Foreign Languages and Literatures* will be held on October 25-26, 1996, at Youngstown State University. Workshops and other presentations on topics related to the teaching of foreign languages at all levels.

Inquires to:

Foreign Language Conference
Department of Foreign Languages
Youngstown State University
Youngstown, Ohio 44555
Phone (216) 742-3461

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (ALL) is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. ALL publishes critical, integrative articles, and research reports on adult language learning for functional purposes. The editor encourages the submission of original manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-P (ALL)
ATTN: Editor (L. Woytak)
Defense Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784.
Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run between 30 and 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author
JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. The abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

The list of references should be submitted on a separate page of the manuscript with the centered heading: *References*. The entries should be arranged alphabetically by surname of authors. The sample list of references below illustrates format for bibliographic entries:

Dulay, H. & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly*, *16* (1), 93-95.
Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

Reference citations in the text of the manuscript should include the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that is being quoted originally appeared, e.g., (Jones, 1982, pp. 235-238). All works cited in the manuscripts must appear in the list of references, and conversely, all works included in the list of references must be cited in the manuscript.

Notes should be used for substantive information only, and they should be numbered serially throughout the manuscript. Subsequently they all should be listed on a separate page title *Notes*.

Specifications for Floppy Disks

Where feasible, manuscripts are preferred on 3.5" disk. Manuscripts produced on one of the DOS-based or Macintosh systems should be formatted as MS-DOS file at double density, if possible.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____

2. Disk is formatted as: double__ high density__

3. Remarks: _____

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

890

* U.S. GOVERNMENT PRINTING OFFICE: 1996-785-252/59074 PIN: 074108-1/N-000

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

Professional Bulletin 65-95-1&2
United States Army
Approved for public release.
Distribution is unlimited.

Applied Language Learning

VOLUME 7 • NUMBERS 1&2
1996

393

Applied Language Learning

Colloquium Proceedings
on

Validity Issues in the Assessment of L2 Learner Strategies

Volume 7 • Numbers 1 & 2

1996

Defense Language Institute
Foreign Language Center
Presidio of Monterey

894

Applied Language Learning
PB 65-96-1 & 2

The mission of Professional Bulletin 65-96, *Applied Language Learning* (US ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

*Applied Language Lea*rning, published semi-annually by the Defense Language Institute, Foreign Language Center (DLIFLC), presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

DENNIS J. REIMER
General, United States Army
Chief of Staff

Official:
JOEL B. HUDSON
Administrative Assistant to the Secretary of the Army
02136

Daniel D. Devlin
Colonel, United States Army
Commandant, DLIFLC

Eugene F. Beauvais
Colonel, United States Air Force
Assistant Commandant, DLIFLC

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the individual authors. *Applied Language Learning* is available on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster:
Send change-of-address information to:

Applied Language Learning
Defense Language Institute, Foreign Language Center
Presidio of Monterey, CA 93944-5006
United States of America

Applied Language Learning
PB-65-96-1 & 2

Lidia Woytak
Editor

Editorial Board, DLIFLC

Eugene F. Beauvais
Colonel
Command Group

Ray T. Clifford
Provost

Christine M. Campbell
Project Director
Testing

Martha Herzog
Dean
Curriculum and Instruction

John L. D. Clark
Dean
Research and Testing

John A. Lett, Jr.
Director
Research and Analysis

Editorial Advisory Board

Kathleen M. Bailey
Professor
Monterey Institute of Interantional Studies

Stephen D. Krashen
Professor
University of Southern California

David P. Benseler
Professor
Case Western Reserve University

Leo Van Lier
Professor
Monterey Institute of International Studies

Special Acknowledgment

The members of the editorial board of *Applied Language Learning express warm* words of appreciation to Professor **June K. Phillips,** the outgoing member, for her contribution to the journal.

Applied Language Learning

1996 Volume 7 Numbers 1 & 2

viii Preface
Lidia Woytak

1 Introduction: Validity Issues in the Assessment of L2 Learner Strategies
Roberta G. Abraham and Roberta J. Vann

Articles

5 Verbal Reports as a Source of Insights into Second Language Learner Strategies
Andrew D. Cohen

25 Employing a Questionnaire to Assess the Use of Language Learning Strategies
Rebecca L. Oxford

47 Validity Issues in Computer-Assisted Strategy Assessment
Carol A. Chapelle

61 Using Task Products to Assess Second Language Learning Processes
Roberta G. Abraham and Roberta J. Vann

General Information

91 Calendar of Events

95 Information for Contributors

Preface

Applied Language Learning occasionally publishes collections of colloquium papers on adult language learning for functional purposes. The 1991 issue of the journal (Vol. 2, No. 2) presented proceedings from an invitational conference on "Improving Foreign Language Teaching through Technology." The conference was held at the Defense Language Institute on October 30, 1990.

This special issue of *Applied Language Learning* contains colloquium proceedings on "Validity Issues in the Assessment of Second-Language Learner Strategies" presented at the American Association for Applied Linguistics (AAAL) Conference in Long Beach, California, in March 1995.

If you have recently organized a colloquium on adult language learning, you may wish to submit those proceedings for publication in *Applied Language Learning*. To submit your proceedings for consideration, send a proposal briefly describing the purpose of the colloquium and the content of individual papers.

Applied Language Learning
Editor, Dr. L. Woytak
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006
United States of America

Lidia T. Woytak
Editor

Applied Language Learning
1996, Vol. 7, Nos. 1 & 2, pp. 1-4

Introduction
Validity Issues in the Assessment of L2 Learner Strategies

Roberta G. Abraham and Roberta J. Vann
Iowa State University

As both producers and consumers of L2 strategy research, we have often wondered about the accuracy of the picture of learner strategies emerging from the literature over the past two decades. During this period a number of methods of assessing learning strategies have been used, beginning with informal observation and "talking with people" reported by Rubin (1975) and including think-alouds (e.g., Cohen & Hosenfeld, 1981), interviews (e.g., O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985), questionnaires (e.g., Politzer, 1983; Oxford, 1986), computer assessment (e.g., Jamieson & Chapelle, 1987) and, more recently, product analysis (Vann & Schmidt, 1993; Abraham & Vann, 1994). Our questions about these methods have included the following:

- Do learners responding to questions in an interview or on a questionnaire really understand what they are being asked?
- Do learners tell the truth in responding to these questions (rather than providing the answer they think the researcher wants to hear)?
- How complete and accurate a picture of the processes are learners able to self-report?
- Why do two different methods of assessing learner strategies sometimes not give the same answers (as in LoCastro, 1994)?
- To what extent do "unnatural" methods like think-alouds reflect the strategies learners normally use in performing a task?
- How accurately and completely do performance data collected on computers or in learner products reflect learner processes?

Thus, when we were invited to organize a colloquium for the 1995 American Association for Applied Linguistics conference, we knew immediately what our topic should be: the validity issues surrounding the assessment of L2 learner strategies. As a basis for exploring these issues, we turned to Messick's (1989) insightful discussion in which he notes that "[v]alidity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (p. 13, italics in original). He points out that it is not the method or observation device per se that is to be validated, but rather "the inferences derived from

test scores or other indicators—inferences about score meaning or interpretation and about the implications for action that the interpretation entails" (p. 13). Thus, in validation, we must look not only at the method of collecting data, but also, and more importantly, at how the data are interpreted, that is, what inferences are drawn from the results and how these inferences are justified, and what uses can legitimately be made of these interpretations. Messick stresses that validity is a "unitary concept," and that it is a matter of degree, not of all or none. Furthermore, over time, the existing validity evidence becomes enhanced (or contravened) by new findings" (p.13). [1]

For the colloquium, we asked three colleagues, each of whom has been associated with a key method for investigating L2 learner strategies, to join us in this discussion: Andrew Cohen, Rebecca Oxford, and Carol Chapelle. We all considered the following questions in our presentations:

- What have we learned about the use of our particular method for obtaining information about learner strategies over the past 15-20 years?
- How have particular users of the method defined the term "strategy"?
- What can one legitimately infer about strategy use from data obtained from the method? Conversely, what inferences cannot or should not be drawn from data obtained from the method?
- In what way has the method been particularly useful (to researchers and practitioners)?
- What problems, perhaps unforeseen, have arisen with the method in assessing strategy use?

The articles in this issue are revised versions of the papers these questions elicited. The first two articles deal with relatively well established methods of eliciting information about learner strategies (verbal reports and questionnaires), while the last two discus methods less frequently associated with strategy research (computer-assisted assessment and analysis of learner products).

In the first article, Andrew Cohen contrasts three types of verbal report methods that have been used to illuminate L2 learner strategies, and then discusses several research practices that should be taken into account in interpreting the data. The last portion of his paper addresses ways in which verbal report methods can become more "robust" and the reporting of research using verbal report more complete, thus permitting more valid interpretation and more appropriate use of the results.

Rebecca Oxford's article on the use of questionnaires to assess the use of L2 learning strategies traces the development of her *Strategy Inventory for Language Learning (SILL)* and provides various types of support for the validity claim that it "measures what it purports to measure." Oxford is careful to note what kind of information questionnaires can and cannot provide, and concludes with a number of uses that can legitimately be made of the SILL.

The third article, by Carol Chapelle, reviews computer-assisted strategy assessment over the past fifteen years. Chapelle then looks at two studies where learner

strategies were inferred from performance data collected by the computer, describing the researchers' justifications for their inferences and suggesting further arguments that could have been made. While there are problems in computer-assisted strategy assessment, Chapelle sees this method as a non-obtrusive means of observing learner behavior that can complement other methods in integrating SLA and pedagogical research.

The final article describes two of our own studies in which learner strategies are inferred from learner products. Building on research in communication strategies that shows how products examined within appropriate conceptual frameworks can reveal learner intentions and processes, we describe evidence for the validity of our interpretations of our products, cloze responses and academic reading notes, as reflections of learner processes. In this article we hope to dispel the notion that process and product are dichotomous and to suggest that products provide yet another means of discovering how learners learn.

Note

[1] According to Messick (1989), this view of validity has evolved from the more traditional one in which three types of validity were identified: content, criterion-related (predictive and concurrent), and construct. He shows how content and criterion-related evidence can provide support for construct validity, which in turn should be complemented by consideration of the implications of using a specific instrument in a specific situation. All of these types of evidence can contribute to the on-going validation process.

References

Abraham, R. G., & Vann, R. J. (1994, March). *Metacognition in the cloze performance of nine Lebanese ESL learners.* Paper presented at the Annual Conference of the American Association for Applied Linguistics, Baltimore.

Cohen, A., & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. *Language Learning, 31*(2), 285-313.

Jamieson, J., & Chapelle, C. (1987). Working styles on computers as evidence of second language learning strategies. *Language Learning, 37*(4), 523-544.

LoCastro, V. (1994). Learning strategies and learning environments. *TESOL Quarterly, 28*(2), 409-414.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement,* 3rd ed. (pp. 13-103). New York: American Council on Education & MacMillan.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. P. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*(1), 21-46.

Oxford, R. L. (1986). *Development and psychometric testing of the Strategy Inventory for Language Learning (SILL).* ARI Technical Report 728. Alexandria, VA: Army Research Institute. Appendices as ARI Research Note 86-92.

Politzer, R. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievement. *Studies in Second Language Acquisition, 6*(1), 54-65.

Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly, 9*(1), 41-51.

Vann, R. J. & Schmidt, H. H. (1993, April). *Using student notes to assess reading strategies*. Paper presented at the 27th Annual Convention and Exposition of Teachers of English to Speakers of Other Languages, Atlanta.

Applied Language Learning
1996, Vol. 7, Nos. 1 & 2 , pp. 5-24

Verbal Reports as a Source of Insights into Second Language Learner Strategies

Andrew D. Cohen
University of Minnesota

Since a number of different definitions of second-language (L2) strategies have appeared in the literature, first I define the different kinds of behavior that fall under the rubric of L2 learner strategies and, afterwards, classify different types of verbal report. Next, I briefly indicate areas in which verbal report has made a contribution to our understanding of L2 learner strategies. Then, I focus on problematic issues regarding the methodology itself. Finally, I note ways to refine verbal report methods as I list the types of needed information to understand the particular verbal report. I hope that this information should assist researchers in comparing across studies and in replicating studies that have appeared in the literature. Hence, I do not end on a justification of verbal report methods as others have done, but rather on the fine-tuning of such methods.

Within the last fifteen years, verbal reports have been used increasingly as a source of data on the strategies of learning a second or foreign language. A major impetus for this research technique in L2 acquisition has been its successful use in first-language studies, especially in research on cognitive processes in first-language reading and writing. In fact, this impetus continues to manifest itself; at least two new books on verbal reports in first language have just appeared, one dealing with reading (Pressley & Afflerbach, 1995) and the other dealing with writing (Smagorinsky, 1994).

Defining "Second-Language Learner Strategies"

Second language learner strategies encompass both *L2 learning* and *L2 use strategies*. Taken together, they constitute the steps or actions selected by learners either to improve the **learning** of an L2, the **use** of it, or both. *Language-use strategies* actually include *retrieval strategies, rehearsal strategies, "cover" strategies,* and *communication strategies*. What makes the definition for language-learning and language-use strategies broad is that it encompasses those actions that are clearly aimed at language learning, as well as those that may well lead to

learning but which do not ostensibly have learning as their primary goal.[1] Whereas *language-learning* strategies have the explicit goal of assisting learners in improving their knowledge in a target language, *language-use* strategies focus primarily on employing the language that learners have in their current interlanguage.

Thus, strategies for learning the subjunctive in Spanish as a foreign language, for example, could include grouping together and then memorizing the list of verbs that take a subjunctive in constructions like *quiero que vengas* ("I want you to come"), or noticing the difference in imperfect subjunctive inflections between the *-ar* conjugation (e.g., *cantara*) and the *-er* and *-ir* conjugations (e.g., *comiera, existiera*). The specific strategies for memorizing this group might involve writing these verbs inside a box in the notebook and reviewing the contents of the box regularly, as well as noting what these verbs have in common semantically. Language-learning strategies would also include strategies for learning new vocabulary, such as using flash cards, possibly with keyword mnemonics to jog the memory if necessary.[2]

Strategies for **using** the subjunctive include four subsets of strategies: *retrieval strategies, rehearsal strategies, "cover" strategies,* and *communication strategies*. In the above example with the subjunctive, *retrieval strategies* would be selected for retrieving the subjunctive forms when the occasion arises in or out of class, and for choosing the appropriate forms. For those learners who keep a list of verbs taking the subjunctive, a strategy may involve visualizing the list and cross-checking to make sure that the verb that they wish to use in the subjunctive form actually requires the subjunctive. Likewise, a language-use strategy would entail using the keyword mnemonic to retrieve the meaning of a given vocabulary word. So, say that a learner encounters the verb *ubicar* (to locate), which she had learned by means of the keyword mnemonic *ubiquitous*, and she wants to retrieve the meaning of the word. The language-using strategies would include any efforts by the learner to retrieve the meaning of the word *ubicar*—involving the linking of the Spanish sounds */ubik/* with the English */yub k/*, and then perhaps seeing an image of someone who keeps turning up everywhere the language learner looks.

Language-use strategies also include strategies for rehearsing target language structures (such as form-focused practice), as well as strategies for covering one's self in the language classroom (such as participating in classroom tasks to look good in front of other students or the teacher, without intending to learn or communicate any particular aspect of the target language).[3] An example of rehearsal would be form-focused practice, for example, practicing the subjunctive forms for different verb conjugations. An example of a "cover" strategy would be using a memorized and not fully-understood phrase in an utterance in a classroom drill in order to keep the action going. Some cover strategies reflect efforts at simplification (e.g., learners use only that part of a phrase that they can deal with), while other such strategies complexify the utterance (e.g., saying something by means of an elaborate and complex circumlocution because the finely-tuned vocabulary is lacking or to avoid using the subjunctive). Both cases represent an

attempt to compensate for gaps in target language knowledge.

Communication strategies constitute a fourth subset of language-use strategies, with the focus on conveying meaningful information that is new to the recipient. Such strategies may or may not have an impact on learning. For example, learners may use a vocabulary item encountered for the first time in a given lesson to communicate a thought, without any intention of trying to learn the word. In contrast, they may insert the new vocabulary item into their communication expressly to promote their learning of it.

It is fair to say that verbal report data have enhanced our understanding of all the above types of L2 learner strategies. For example, verbal report data have been collected from learners as they generate mnemonic devices for remembering new vocabulary words and as they attempt to retrieve these words later by means of these mnemonic devices (see Cohen & Aphek, 1980, 1981). Likewise, verbal report data have been used to reveal instances where learners are using material over which they have little control. For example, an earlier study by Cohen and Aphek (1979) revealed an instance of a survival-oriented cover strategy, when an adult learner of Hebrew used a nonexistent form *yariya* instead of *ered* (I will get off). Retrospective verbal report provided by the learner in class revealed that although she had been exposed to the future, she had not as yet achieved productive control over the tense. She explained that she lived on a street called *Yordei Hasira* and knew that the first word of the street name was also derived from the verb *laredet* 'to get off.' So she improvised by making up a word that might convey the future tense. This insight was only made possible through the use of verbal report—in this case, through immediate retrospection.

Since verbal report as it has been applied to the field of learner strategies is not one measure, but rather encompasses a variety of measures intended to provide mentalistic data regarding cognitive processing, the next section provides a brief classification of the report types.

Classifying Verbal Reports

Verbal reports include data that reflect (1) *self-report*: learners' descriptions of what they do, characterized by generalized statements about learning behavior—e.g., "I tend to be a speed listener" (2) *self-observation*: the inspection of specific rather than generalized language behavior, either introspectively, i.e., within 20 seconds of the mental event, or retrospectively—e.g., "What I just did was to skim through the incoming oral text as I listened, picking out key words and phrases" and (3) *self-revelation*: "think-aloud," stream-of-consciousness disclosure of thought processes while the information is being attended to—e.g., "Who does the 'they' refer to here?" Verbal reports can and usually do comprise some combination of these (Radford, 1974; Cohen & Hosenfeld, 1981; Cohen, 1987). Self-report data tend to appear frequently on questionnaires that ask learners to describe the way they usually learn and use language. Self-observation implies reference to

some actual instance(s) of language learning or use. For example, entries in journals or diaries that retrospectively describe some language-learning or language-use event involving the subjunctive would count as retrospective self-observation. Self-revelation or think-aloud data are only available at the time that the language learning or use events are taking place, and imply that the respondent is describing, for example, the struggle to use the correct form of the subject, and not attempting to analyze this struggle. Thoughts which are immediately analyzed would constitute introspective self-observation—for example, "Now, does this utterance call for the present or imperfect subjunctive? Let me see..."

Examples of learner-strategy studies containing verbal reports in the form of self-report interviews and questionnaires include Naiman, Fröhlich, Stern, and Todesco (1978); O'Malley, Chamot, Stewner-Manzanares, Kupper, and Russo (1985); Wenden (1985); Ramírez (1986); and Oxford, Nyikos, and Crookall (1987). In such studies, the respondents answered interview questions or completed written questionnaires about their language strategies. Since self-report has been shown to be somewhat removed from the cognitive events being described, this approach may produce data of questionable validity. Questionnaire items are more likely to elicit learners' **beliefs** about what they do, rather than what they actually do. Efforts are often made by investigators to increase the extent of self-observational and self-revelational data and to decrease the amount of self-report. The purpose is to obtain data that describe the learning event at or near the moment it occurs. Such data might be expected to reflect accurately what learners actually do than might the response to a questionnaire item calling for a description of generalized behavior.

Contributions of Verbal Report to Understanding Learner Strategies

Despite frequent criticism (Seliger, 1983; Afflerbach & Johnston, 1984; Olson, Duffy, & Mack, 1984; Lyons, 1986), verbal report has gained popularity in the last several decades because it provides data on cognitive processes and learner responses that otherwise would have to be investigated only indirectly.[4] Furthermore, verbal report has at times provided access to the reasoning processes underlying cognition, response, and decision making. We note that the use of verbal report protocols in L2 learning-strategy investigations has benefited greatly from the extensive use of this research methodology in the native language. Such work, especially in reading and writing (e.g., Garner, 1982; Flower & Hayes, 1984), has paved the way for much of the L2 work.

A recent book by Pressley and Afflerbach (1995) focuses on the use of verbal reports of first-language reading, once again constituting an excellent compendium of ideas for L2 researchers. The authors refer to verbal reports as "a maturing methodology with much interesting work already accomplished and considerable work to be done" (p.1). They demonstrate how the use of verbal report (whether as an exploratory methodology or as a means for testing hypotheses about

reading) has yielded an elegant description of reading. They provide a detailed description of what they refer to as *before reading, during reading, after reading, monitoring* and *evaluating strategies*, based on a review of 38 primary-data studies. As the authors put it, "The think-alouds were extremely revealing about the dynamics of comprehension difficulties and how understandings of text shift in reaction to comprehension difficulties and surprises in text" (p. 38).

With regard to L2 learning and use, verbal report methods—primarily reflecting self-revelation and self-observation—have been employed as a means of describing strategies in the learning and use of L2 vocabulary (e.g., Cohen & Aphek, 1978, 1981; Neubach & Cohen, 1988; Chern, 1993, Huckin & Bloch, 1993), in L2 listening (e.g., Murphy, 1987), in L2 speaking (e.g., Robinson, 1991, Cohen & Olshtain, 1993; Cohen, Weaver, & Li, 1995), in L2 reading (e.g., Hosenfeld, 1984; Block, 1986; Cavalcanti, 1987; Kern, 1994), and in L2 writing (e.g., Zamel, 1983; Raimes, 1987; Cohen & Cavalcanti, 1987, 1990; Skibniewski, 1990). Verbal report is also used for investigating the subset of L2 communication strategies, especially those used in compensating for gaps in communicative ability (e.g., Poulisse, Bongaerts, & Kellerman, 1986; Poulisse, 1989). In addition, verbal report is used with tasks that combine most or all of the strategy areas, such as in investigating the strategies used in translation of texts (Faerch & Kasper, 1986; Borsch, 1986; Gerloff, 1987; Krings, 1987) and those used in taking L2 tests (Cohen, 1984, 1994a, 1994b; Stemmer, 1991; Gordon, 1987; Anderson, 1991; Nevo, 1989).

Despite the extensive use of verbal report methods in numerous recent studies, readers are still sometimes uncertain as to the inferences that they can legitimately make on the basis of these reports. At the same time that Pressley and Afflerbach (1995) refer to verbal reports as a maturing method, they also rightly refer to it as an "underdeveloped" one (p.119). For this reason, I will now consider a series of problematic areas regarding the methodology, with an eye to where development needs to take place.

Verbal Report Methodology

Immediacy of Verbal Report

A distinction has been made in the literature between self-revelational data in the form of immediate, on-line think-aloud protocols (which involve no editing or analysis), on the one hand, and self-observational data in the form of introspective or retrospective self-observation on the other. Ericsson and Simon (1993) have advocated the collection of self-revelational data over other approaches to verbal report because asking questions **only** about what was heeded in short-term memory was seen as a means of making such reports more reliable in that there is no strain on the memory to reconstruct past thoughts.[5] In sharp contrast to this methodological position, the Pressley and Afflerbach (1995) survey of studies in L1 reading found considerable variation as to the immediacy of the reporting and

the amount of interpretation respondents were asked to provide (p. 22).

The researchers found not only self-revelational protocols but also self-observational reports that were collected after each sentence, after each episode, at signaled spots in the text (usually two or more sentences), after every two minutes, at the end of the text, or whenever the readers wanted. Thus, there was a fluctuation both within and across studies as to whether subjects were asked to provide think-aloud, introspective (i.e., within 20 seconds of the event), or retrospective reports (separated somewhat in time from the actual reading). Pressley and Afflerbach (1995) give one explanation for this departure from exclusive use of the think-aloud approach—namely, that to obtain verbal report of otherwise automatized cognition, there is a need to slow down the process by using, for example, the interruptive methods listed above (p. 9).

Not only did Pressley and Afflerbach (1995) have difficulty in determining if verbal reports in the studies that they reviewed reflected traces remaining in short-term memory or rather the subjects' reconstructions of what happened as they read. They were also unable to determine whether there was substantive difference in quality between think-aloud data produced when subjects performed no analysis and the self-observational data when they analyzed what they were thinking (p. 128). The reasons they gave for their inability to make a comparison were (1) there was too little systematic study of this issue in the given research reports, and (2) the verbal reporting itself was influenced differentially by the nature of the training, coaching, or prompting that the respondents received before and during the reporting phase.

Greene and Higgins (1994) considered the issue of delay in the case of retrospective verbal report after the completion of a writing task. The investigators offered four suggestions for improving the reliability and validity of such data: (1) minimizing the time between the process and report by obtaining a report immediately after a writer completes a task, (2) designing prompts that can help writers better access detailed information from their short- and long-term memory (e.g., through the use of concrete examples and contextual cues), (3) making clear to the respondents the purpose of the retrospective accounts, and (4) reporting one's findings in ways that enable readers to see how the conclusions have been derived from the data (e.g., by including enough data in a report so that readers can make their own assessments about the value of research based on retrospection).[6]

Respondents' Role in Interpreting the Data

There are researchers who are wary about having subjects interpret why they are doing something. Their rationale is that a request to provide interpretation is more likely to influence how the respondents perform continuing phases of the same task. In addition, they see the asking of a "why" question as likely to produce unreliable answers if at the time the respondent is not thinking about why he/she is doing the action (Ericsson & Simon, 1993, p. 7).[7] Hence, they recommended that

interpretation of verbal report be left to researchers, rather than, asking the respondents to categorize their cognitions. Despite these recommendations, Pressley and Afflerbach's (1995) review of 38 primary data studies of L1 reading found that many studies went beyond having readers simply report their thoughts, and requested them to interpret their processes as well (p. 21). Presumably, the insights from self-observation offer a rich enough source of information not available through think-aloud protocols alone that researchers are willing to risk threats to the reliability of the verbal report tasks in order to obtain the data.

Prompting for Specifics in Verbal Report

Early descriptions of verbal report methods usually included the stipulation that respondents not be given instructions as to what to report on. They were to be left to their own devices since any instructions might lead to biased processing. But anyone who has been faced with analyzing pages of transcribed **undirected** verbal report protocols has seen that such data are likely to be general and incomplete. So, even methodological hard-liners like Ericsson and Simon (1993) favor instructions to the respondents to make the verbal reports complete (p. 11).

Thus many studies now do include instructions to elicit particular cognitive behaviors. For example, reading researchers have cued different processes in the different studies. Pressley and Afflerbach (1995) found one study that requested that subjects create a summary of what they read, and in which the respondents were informed about the importance of summarization, a second that asked respondents to attend to content and style when reading; and others that required subjects to draw inferences. The authors conclude that prompting respondents to use particular processes may be necessary: "it is reasonable to prompt [processes] in order to assure that a sample of the target processes will, in fact, be observed" (p. 133). With regard to post-experimental assessment, Cantor, Andreassen, and Waters (1985) have found that more valid information is produced if the cues involve specific items from the experiment (in their case, animal episodes and geometric form episodes).

Guidance in Providing Verbal Reports

Not only has it proven effective to have respondents receive specific prompts as to what to report about, but it has also been seen that instruction in how to provide verbal report for a given task improves the quality of the data. Ericsson and Simon (1993) have found that to assure that the verbal report does not interfere with the task at hand, there must be warm-up trials after the instructions with tasks that yield easy-to-analyze think-aloud, introspective, and retrospective reports. The researchers suggest that—to ensure consistency—subjects be given trials on these warm-up tasks until they are able to make verbal reports without confounding them with explanations and justifications (p. xxxii). "In some studies, more

extensive warm-up procedures are used explicitly to *train* the subjects to conform to the think-aloud instructions" (Ericsson & Simon, 1993, p. 82). In a study in which subjects were asked not only to think aloud, but also to give a reason for each response they made before keyboarding it into the computer, the respondents who provided verbal report after receiving training improved more on the computerized cognitive task than those who did not receive the training (Berry & Broadbent, 1984). In the review of 38 primary studies of verbal report in L1 reading, Pressley and Afflerbach (1995) found that while in some studies the respondents were given an opportunity to practice, in others they were not (p. 22).

Reactive Effects of Verbal Report

Verbal report that involves intervening during the performance of a task has been criticized for the inevitable reactive effects that such intervention causes. Stratman and Hamp-Lyons (1994), for example, conducted an exploratory study to determine the extent of reactivity, in which they had writers engage in two revision tasks eight weeks apart, one with think-aloud verbal reports. All subjects were trained in providing think-aloud protocols. The researchers found for the eight subjects in their study that thinking aloud increased the number of new "word-level" errors (morphological, tense, and spelling, p. 103). Contrary to the investigators' expectations, thinking aloud was found to inhibit word or phrase additions. They also found that while thinking aloud did not have an impact on complex meaning changes at the microstructural level, it stimulated the production of entirely new sentences (p. 107). They concluded that thinking aloud **does** alter the nature of processing in the revision phase of writing. They posited that think-aloud protocols may systematically influence the correction of organizational-level errors (i.e., reordering of displaced sentences, adjusting faulty paragraph boundaries, detection of faulty pronoun references, detection of redundancies, detection of word-level errors—in morphology, tense, and spelling—and introduction of new word-level errors) and influence the amount and kind of microstructural meaning changes as well.

While the 1994 study by Stratman and Hamp-Lyons on the use of verbal report during the revision phase of writing produced reactive results of a negative nature, a series of other studies would suggest that there may be positive consequences of verbal report. Collecting retrospections (termed *intervention protocols*) at various points during the writing has also been found to improve the reliability of the data collection task (Swanson-Owens & Newell, 1994). It was found that the interruption of writing for the purpose of reflecting on process served as a supportive measure in helping writers learn about composing, and thus to provide scaffolding for a subject's learning during data collection. Similarly positive outcomes of verbal report have been reported for studies in the areas of vocabulary learning and reading as well. For example, Crutcher (1990) conducted a study of vocabulary learning with keywords and obtained retrospective reports for half of

the items. He found that retention of the words was better for those items.

With regard to verbal reports in L2 reading, Nyhus (1994) looked at the attitudes of NNS of English toward the use of verbal report to elicit their reading comprehension strategies. The respondents were seven third-quarter students in the Commanding English Program in General College at the University of Minnesota—a bridge program for refugee and immigrant non-native speakers of English. Five of the respondents were Vietnamese, one Chinese, and one Russian. Most had been in the U.S. for only two to three years. The study looked at their attitudes toward the effects of think-aloud and retrospective verbal report on their reading. They were also asked to assess verbal report as a research methodology.

The respondents were shown a videotape of the researcher reading aloud and providing a think-aloud verbal report from a sociology text. Three excerpts from a sociology text were chosen for use with the respondents. Two were for practice readings and the third for the data collection. Red dots were placed between sentences to remind the respondents to verbalize their thoughts. Two sets of interview questions were developed, the first twelve questions to be asked following the respondents' initial think-aloud verbal report and the second eleven questions to be asked following the respondents' retrospective verbal report. The respondents were asked to read the text as they normally would but to say all of their thoughts aloud, in English. They were told they could read the text silently, but all chose to read it aloud. The respondent and the researcher then listened to the recording of the verbal report and the respondents provided a retrospective verbal report by pausing the tape when they wanted to make additional comments about thoughts that had occurred to them while reading the text. The researcher also had the respondents report on what they had been thinking but not verbalizing. Next, the researcher interviewed the respondents regarding their views about the think-aloud methodology. Then, there was a second interview to elicit attitudes toward the retrospective methodology after the task had been completed.

For the most part, the respondents viewed the effects they attributed to verbal report as beneficial. Most felt that think-aloud verbal report affected their thinking about reading in a positive way. They reported that it enhanced their awareness and assessment of various aspects of the reading process, including an awareness of themselves as readers and of their interaction with the given text. Only two of the seven had negative comments about verbal report, and these were the students whose English was the most limited. Since all verbal report was conducted in English, performing the verbal report in English was most likely to the detriment of those with more limited English skills. There may, in fact, be a second-language threshold below which attempts to provide verbal report in the target language are counterproductive.

Despite several cases of difficulty in reporting in English, all respondents viewed verbal report as useful in various ways. They saw it as a means for placing students at a given level, as a diagnostic tool for determining their specific reading needs at a given level, and as a solitary or group study technique. The students

reported that in small groups they discovered alternative ways of thinking about a text. Retrospective verbal report generated by having readers listen to and comment on a playback of their think-aloud verbal report provided still more insights. It was seen as a means of helping readers, instructors, and researchers alike to gain further insight into readers' thinking and reading processes.

Towards Robust Verbal Report Methods and Complete Write-Ups

What has emerged from this discussion of methodological issues in verbal report as applied to language learner strategies is that we are in need of both more refined measures and more details about the verbal report methods of each study. This more detailed information would facilitate cross-study comparisons regarding both learner strategies and the research methodology itself. So, for example, Pressley and Afflerbach (1995) propose a study of reading strategies that would call for a carefully detailed comparison between think-aloud verbal reports and delayed reports. The study would assess the extent to which ongoing verbal report might interfere with the natural reading processes, and the extent to which delayed stopping after every sentence or few sentences might shift the nature of subsequent reading, if at all. Pressley and Afflerbach (1995) would also wish to investigate the question of how long reports can be delayed before they decay (p. 13). In making their plea for greater completeness in the description of verbal report methods, they include a listing of variables for which more complete and systematic information is desirable (pp. 120-123).

Issues of Method and of Write-Ups for Verbal Reports

Let us now relate Pressley and Afflerbach's listing of variables to second-language studies. The following list includes areas for refining verbal report methods and for encouraging write-ups describing the methods in detail sufficient to ensure comparison across studies.

Subjects' Characteristics

For the purpose of comparison across studies, the educational background of the respondents, their knowledge of the task at hand, and their motivation to perform the task should be made clear. In addition, their level of language proficiency (especially in the case of L2 studies) and their age should be indicated. Pressley and Afflerbach (1995) also suggest that their short-term memory capacity and their spatial ability be noted, but this would entail special psychological testing that is usually not conducted in L2 acquisition research. These authors also stress the need for studies with larger numbers of subjects, since most studies are of individual cases or small groups. Their point is that while the accumulation of small-scale studies of verbal report does help to generate a large-scale picture, comparison across them can be somewhat problematic, especially if the approaches

to data collection are different. The problem is that most researchers do not have the budget to conduct verbal report work with large groups.

Whereas Pressley and Afflerbach (1995) limit themselves to respondents who were performing a task in their native language and providing verbal report in that language, research into L2 learner strategies is faced with the issue of choice of language for verbal reporting. When dealing with groups of speakers of numerous languages, the verbal report protocols may need to be in the target language. In cases where the respondents share the same native language or speak a limited number of languages, it may be advisable to give them a choice as to language of verbal report, since the less proficient they are in the target language, the more difficulty they may experience trying to perform the task and provide verbal report in the target language at the same time. The study by Nyhus (1994), in fact, found that the two poorer ESL readers were the ones reporting difficulty providing the verbal report, which was in the L2. Regardless of whether bilinguals use one language or the other for their verbal report, it is important that the researcher indicate the extent to which one or the other is used.

Characteristics of the Materials

When textual material serves as a stimulus for verbal report data, it would be helpful if the investigator specified the genre of the material, its topic, its length, and its difficulty level for the given respondents. While some or most of these variables may be provided as a matter of course (especially if texts are included in the appendix of the study), Pressley and Afflerbach would request that investigators indicate the fit between the task and the characteristics of the given respondents. Any such details could help other researchers to interpret the findings with greater ease, as well as to attempt replication of the study, if so desired. Perhaps more so in foreign than in native language reading, the genre of the text can make a big difference in the ease of reading. Even if the readers feel comfortable with the genre (e.g., journalistic writing), still they may have difficulty with the specific topic transmitted by means of that genre (e.g., an account of a holiday with which the reader is completely unfamiliar).

Criterion Task

It is imperative for the purpose of comparison that the researcher provide a clear indication of the tasks that the respondents were asked to perform (e.g., in reading research, whether it was free recall, recognition, question answering, summarization, or some combination of these), plus the directions given to the subjects. Pressley and Afflerbach found in the studies they reviewed that the instructions were either not provided or that reference to them was vague. The reason that the instructions are considered so crucial in verbal report work is expressly because of the orienting that takes place through instructions. It is also important to have a clear description of any technical equipment employed in the study (e.g., a multimedia program on CD-ROM). Likewise the goals of the language task should be clear,

as well as the modalities utilized.

Guidance in Verbal Reporting

It is valuable both for purpose of comparison across studies and for replication that information be given as to the nature and extent of guidance that the subjects received in verbal reporting. It is useful to know, for example, whether the subjects received feedback in practice sessions, whether they were coached during the data collection sessions, and if so, the length of the guidance—for example, until they got the behavior correct or until they acted as they were supposed to act. It has become more common to instruct respondents in how to provide verbal report, as well as to coach them as they are providing it (e.g., requesting that they not report on the basis of what they **usually** do, but rather that they stick to what they are actually doing in the given instance).

Methods of Analysis

To help other researchers interpret the findings, it may prove beneficial to include details concerning the development of categories and coding of verbal reports. Further it may be beneficial to include the codes and symbols used in the transcriptions of the verbal report protocols as well—for example, symbols for suprasegmental features, such as tone of voice.[8] Pressley and Afflerbach found that the reporting of these methods was usually incomplete.

Categories Used to Score Verbal Report Protocols

It is helpful for researchers to indicate how the scoring of verbal report protocols is done, since so much interpretive work is involved. If the respondents themselves listen to their verbal reports in order to assist in the interpretation of protocols, as in the case of the study by Nyhus (1994) on the effects of verbal report on L2 reading, it would be important to highlight this feature and describe it fully in the write up phase. Such a procedure has the value of improving the validity of the measure, since the respondents themselves are verifying the accuracy of what they reported (choice of words, completeness of the report, etc.) and possibly adding what they had neglected to mention the first time around. It might even pay to have researchers provide verbal report while they are engaged in the task of making their decisions about how to score given instances of behavior appearing in the protocols. Verbal report protocols of raters of L2, for example, reveal instances where the raters do not understand the categories that they are supposed to be using in their ratings (e.g., "style," "register," and so forth).

Inter-Rater Reliability Checks

In cases where two or more investigators score the data, it would be advisable to run inter-rater reliability checks to determine the extent to which the investigators are using similar criteria in arriving at scores. Information about such checks should be provided in the research report.

Selection of Verbal Report Excerpts for Inclusion in Research Reports

A somewhat subtle issue is that of how the data are chosen for inclusion in reports. Other researchers would want to know how representative such excerpts are of the data set as a whole. There is a concern that the investigators may slant the findings according to the excerpts from the data that they choose to select for inclusion in any reports that they write. It is for this reason that Greene and Higgins (1994) go to some lengths to demonstrate how to represent verbal report data in an equitable way in their study of retrospective verbal report of L1 writing processes.

Theories Used in Framing Verbal Report Study

The researchers are asked to identify the theoretical principles that the verbal report techniques were intended to investigate. Pressley and Afflerbach consider it the researchers' responsibility to provide information as to whether the verbal report measures really reflect the cognitive processes that are reported. This information is necessary to validate the verbal report measures of the study. They contend that the researchers should indicate the relationship between the verbal report and the performance outcomes, much as they do in their own book, by demonstrating that theoretical models of reading (e.g., Baker & Brown, Anderson & Pearson, Dijk and Kintsch, and their own models of constructively responsive reading) are supported by verbal report data obtained from reading studies. As Pressley and Afflerbach (1995) put it,

> As validation efforts proceed, we urge careful attention to the establishment of clear linkages between theory, verbal process reports, and other measures that can be complementary to verbal self-reports. We believe this work will do much to bring verbal reports from the status of a 'bootstrap operation' (Ericsson & Simon, 1993) to a maturing methodology (p. 126).

Most published studies of second language acquisition include a statement of the research questions and the rationale for each one. If the verbal report measures are simply aimed at exploring some aspect(s) of these research questions, then the theoretical underpinnings are probably provided. It is possible, however, that the theoretical rationale for a given verbal report procedure is not overtly clear to the reader of the report. In such cases, the request would be to provide this rationale.

The Validity of Verbal Reports

While the above discussion of nine issues focused mostly on the reliability of the verbal report measures, their validity also comes into play in each and every issue. While larger samples help to make the results more valid, an alternative to increasing the sample size would be to amass a series of well planned and executed small-scale studies. As for the role played by the materials and the tasks in the determination of validity, it is imperative that the consumers of the research results have adequate information about the nature of the materials and about the specific instructions that the respondents were given for performing the task. Such information is crucial in interpreting the verbal report responses received. By the same token, the consumers of the reports need to know the extent to which the respondents were coached on how to perform the task.

Once the data are collected, the analysis procedures also have direct impact on whether the data measure what they purport to measure—that is to say, the rationale for the construction of the analysis categories and then the actual process of data analysis. Did the raters understand and properly use all of the rating categories? With regard to inter-rater reliability (if there is more than one rater), a low correlation would call into question not only the reliability of the ratings but their validity as well.

Furthermore, there is the issue of whether the reported data are comprehensive or selective, and if selective, what this says about the validity of the reporting process. Finally, there is concern that the study not use verbal report simply for its own sake, but rather because the data collection method does, in fact, help to gather data bearing on the theoretical issue(s) at hand.

Summary and Conclusion

The article started by defining and then illustrating the important split between language-learning strategies on the one hand and language use strategies on the other. It then contrasted the three forms of verbal report—self-report, self-observation, and self-revelation—and briefly indicated the contribution that verbal report methods have made to the understanding of language-learning and -use strategies. It then focused on concerns about the appropriate use of these measures and about the nature of reports that include the findings from the use of such measures. The issues included the immediacy of the verbal reporting, the respondents' role in interpreting the data, prompting for specifics in verbal report, guidance in verbal reporting, and the reactive effects of verbal reporting.

The lengthy focus on both refining verbal report methods and on improving the write up of verbal report procedures was intended to underscore the importance of being rigorous both in design and in description. The purpose would be not only to improve the data, but also to assist others in understanding fully what was done, in being able to make comparisons to other studies, and in being able to

replicate the studies. In addition, the point was made that care in the write up can help to dispel arguments that such methodological approaches are not adequately rigorous.

While previous studies have tended to focus on justifying verbal report in the face of criticism from those opposed to it, this article has instead focused on the fine-tuning of verbal report methods. Since by now so many studies using verbal report techniques have emerged, the time has come to provide greater systematicity both in the collection of such data and in the reporting of such studies through the research literature. This article has intended to help researchers ask and get answers to more finely-tuned questions, so that the already valuable findings of verbal report studies might be enhanced by the extra methodological rigor.

Notes

[1] See Cook (1993, Ch. 6), Ellis (1994, Ch. 12), Towel & Hawkins (1994, Ch. 13), and McDonough (1995) for recent reviews of the learning and communication strategy literature, and for discussion of its terminology.

[2] A keyword mnemonic is a word or phrase, usually in the native language of the learner, that is linked by similar sounds to the word to be learned in the target language. The learner then creates an interacting image between this keyword and the target word.

[3] I am grateful to Tim McNamara for suggesting the term "cover" strategies (Personal Communication, July 9, 1996).

[4] While critics have often referred to verbal report data as too qualitative in nature, Hillocks (1994) argues that quantitative studies, while taking the stance of being dispassionate and objective, inherently involve biased interpretations. By the same token, verbal report often relies on **counting** instances of activity in order to arrive at conclusions. Hence, he would argue against categorically labeling verbal report data as qualitative.

[5] The Ericsson and Simon book was originally written in 1984 and was reissued intact in 1993 with a 53-page preface, intended to update the book. The 1984 volume has served for many as **the** authority on how verbal reports are supposed to be conducted. The Pressley and Afflerbach volume constitutes perhaps the first effort to determine the fit between Ericsson and Simon's methodological recommendations and actual uses made of the methodology in the field.

[6] Current research policies at many institutions now require that respondents be fully informed as to what they will be asked to do and that they give their written consent. So in essence, the days of concealing the true motives from the respon-

dents are waning. Furthermore, it may be counterproductive for the purposes of the study to have the subjects distracted for even a portion of the time by anxieties concerning the uses to be made of their responses.

[7] Actually both reliability and validity are of concern here. First, there is the concern that the measure produce data that are consistent within a given verbal report session and across sessions of a similar nature. The second concern is that the data be valid—i.e., that they actually constitute examples of what they purport to be. Hence, reliability is a contributing factor in the determination of validity.

[8] Of course, verbal report data may also be collected in the written modality, as has been done in various studies (e.g., Robinson 1991). In such cases, there would be no suprasegmentals.

References

Afflerbach, P., & Johnston, P. (1984). On the use of verbal reports in reading research. *Journal of Reading Behavior, 16*(4), 307-322.

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal, 75*(4), 460-472.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology, 36A*, 209-231.

Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly, 20*(3), 463-494.

Borsch, S. (1986). Introspective methods in research on interlingual and intercultural communication. In J. House & S. Blum-Kulka (Eds.), *Interlingual and intercultural communication*. Tuebingen: Gunter Narr, 195-209.

Cantor, D. S., Andreassen, C., & Waters, H. S. (1985). Organization in visual episodic memory: Relationships between verbalized knowledge, strategy use, and performance. *Journal of Experimental Child Psychology, 40*, 218-232.

Cavalcanti, M.C. (1987). Investigating FL reading performance through pause protocols. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research*. Clevedon, England: Multilingual Matters, 230-250.

Chern, C-L. (1993). Chinese students' word-solving strategies in reading in English. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning*. Norwood, NJ: Ablex, 67-85.

Cohen, A. D. (1984). On taking language tests: What the students report. *Language Testing, 1*(1), 70-81.

Cohen, A. D. (1987). Studying language learning strategies: How do we get the information? In A. L. Wenden & J. Rubin (Eds.), *Learner strategies in language learning*. Englewood Cliffs, NJ: Prentice-Hall International, 31-40.

Cohen, A. D. (1990). *Language learning: Insights for learners, teachers, and*

researchers. NY: Newbury House/Harper & Row.

Cohen, A. D. (1994a). *Assessing language ability in the classroom*. Boston: Newbury House/Heinle & Heinle.

Cohen, A. D. (1994b). English for academic purposes in Brazil: The use of summary tasks. In C. Hill & K. Parry (Eds.), *From testing to assessment: English as an international language*. London: Longman, 174-204.

Cohen, A. D., & Aphek, E. (1979). *Easifying second language learning*. A research report under the auspices of Brandeis University and submitted to the Jacob Hiatt Institute, Jerusalem. Educational Resources Information Center, ERIC ED 163 753.

Cohen, A. D., & Aphek, E. (1980). Retention of second-language vocabulary over time: Investigating the role of mnemonic associations. *System, 8*(3), 221-235.

Cohen, A. D., & Aphek, E. (1981). Easifying second language learning. *Studies in Second Language Acquisition, 3*(2), 221-235.

Cohen, A. D., & Cavalcanti, M. (1987). Giving and getting feedback on compositions: A comparison of teacher and student verbal report. *Evaluation and Research in Education, 1*(2), 63-73.

Cohen, A. D., & Cavalcanti, M.C. (1990). Feedback on compositions: Teacher and student verbal reports. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 155-177). Cambridge: Cambridge University Press.

Cohen, A. D., & Hosenfeld, C. (1981). Some uses of mentalistic data in second-language research. *Language Learning, 31*(2), 285-313.

Cohen, A. D., & Olshtain, E. (1993). The production of speech acts by EFL learners. *TESOL Quarterly, 27*(1), 33-56.

Cohen, A. D., Weaver, S. J., & Li, T-Y. (1995). The impact of strategies-based instruction on speaking a foreign language. Unpublished research report. Minneapolis: Center for Advanced Research on Language Acquisition, University of Minnesota.

Cook, V. (1993). *Linguistics and second language acquisition*. New York: St. Martin's Press.

Crutcher, R. J. (1990). The role of mediation in knowledge acquisition and retention: Learning foreign vocabulary using the keyword method. Tech Report No. 90-10. Boulder: University of Colorado, Institute of Cognitive Science.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Revised Edition. Cambridge, MA: MIT Press.

Faerch, C., & Kasper, G. (1986). One learner--two languages: Investigating types of interlanguage knowledge. In J. House & S. Blum-Kulka (Eds.), *Interlingual and intercultural communication* (pp. 211-227). Tuebingen: Gunter Narr.

Flower, L., & Hayes, J. R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication, 1*(1), 120-160.

Garner, R. (1982). Verbal-report data on reading strategies. *Journal of Reading Behavior, 14*(2), 159-167.

Gerloff, P. (1987). Identifying the unit of analysis in translation: Some uses of think-aloud protocol data. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 135-158). Clevedon, UK: Multilingual Matters.

Gordon, C. (1987). The effect of testing method on achievement in reading comprehension tests in English as a foreign language. Unpublished M.A. Thesis, Tel-Aviv University, Ramat-Aviv, Israel.

Greene, S., & Higgins, L. (1994). "Once upon a time:" The use of retrospective accounts in building theory in composition. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 115-140). Thousand Oaks, CA: Sage.

Hillocks, G., Jr. (1994). Interpreting and counting: Objectivity in discourse analysis. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 185-204). Thousand Oaks, CA: Sage.

Hosenfeld, C. (1984). Case studies of ninth grade readers. In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a foreign language* (pp. 231-249). London: Longman.

Huckin, T., & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153-178). Norwood, NJ: Ablex.

Kern, R. G. (1994). The role of mental translation in L2 reading. *Studies in Second Language Acquisition, 16*(4), 441-461.

Krings, H. P. (1987). The use of introspective data in translation. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp.159-176). Clevedon, UK: Multilingual Matters.

Lyons, W. (1986). *The disappearance of introspection.* Cambridge, MA: MIT Press.

McDonough, S. (1995). *Strategy and skill in learning a foreign language.* London: Edward Arnold.

Murphy, J. M. (1987). The listening strategies of English as a second language college students. *Research and Teaching in Developmental Education, 4*(1), 27-46.

Naiman, N., Fröhlich, M., Stern, H., & Todesco, A. (1978). *The good language learner.* Research in Education Series, No. 7. Toronto: Ontario Institute for Studies in Education.

Neubach, A., & Cohen, A. D. (1988). Processing strategies and problems encountered in the use of dictionaries. *Dictionaries: Journal of the Dictionary Society of North America, 10*, 1-19.

Nevo, N. (1989). Test-taking strategies on a multiple-choice test of reading comprehension. *Language Testing, 6*(2), 199-215.

Nyhus, S. E. (1994). Attitudes of non-native speakers of English toward the use of verbal report to elicit their reading comprehension strategies. Plan B Paper,

M.A. in English as a Second Language. Minneapolis, MN: University of Minnesota.

Olson, G.M., Duffy, S.A., & Mack, R.L. (1984). Thinking-out-loud as a method for studying real-time comprehension processes. In D.E. Kieras et al. (Eds.), *New methods in reading comprehension research* (pp.253-286). Hillsdale, NJ: Lawrence Erlbaum.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*, 21-46.

Oxford, R., Nyikos, M., & Crookall, D. (1987). *Learning strategies of university foreign language students: A large-scale study*. Washington, D.C.: Center for Applied Linguistics.

Poulisse, N. (1990). *The use of compensatory strategies by Dutch learners of English.* Dordrecht: Foris/Berlin: Mouton De Gruyter.

Poulisse, N., Bongaerts, T., & Kellerman, E. (1986). The use of retrospective verbal reports in the analysis of compensatory strategies. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research.* Clevedon, UK: Multilingual Matters, 213-229.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum.

Radford, J. (1974). Reflections on introspection. *American Psychologist, 29*(4), 245-250.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning, 37*(3), 439-467.

Ramirez, A.G. (1986). Language learning strategies used by adolescents studying French in New York schools. *Foreign Language Annals, 19*(2), 131-141.

Robinson, M. (1991). Introspective methodology in interlanguage pragmatics research. In G. Kasper (Ed.), *Pragmatics of Japanese as native and target language*. Honolulu, HI: Second Language Teaching and Curriculum Center, University of Hawaii (Technical Report; v. #3), 29-84.

Seliger, H.W. (1983). The language learner as linguist: Of metaphors and realities. *Applied Linguistics, 4*(3), 179-191.

Skibniewski, L. (1990). The writing processes of advanced foreign language learners: Evidence from thinking aloud and behavior protocols. In J. Fisiak (Ed.), *Papers and studies in contrastive linguistics* (pp. 193-202). Poznan: Adam Mickiewicz University.

Smagorinsky, P. (Ed.). (1994). *Speaking about writing: Reflections on research methodology*. Thousand Oaks: Sage Publications.

Stemmer, B. (1991). *What's on a C-test taker's mind? Mental processes in C-test taking*. Bochum: Universitatsverlag, Dr. N. Brockmeyer.

Stratman, J. F., & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud protocols. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on*

research methodology (pp. 89-112). Thousand Oaks, CA: Sage.

Swanson-Owens, D., & Newell, G. E. (1994). Using intervention protocols to study the effects of instructional scaffolding on writing and learning. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 141-162). Thousand Oaks, CA: Sage.

Towell, R., & Hawkins, R. (1994). *Approaches to second language acquisition.* Clevedon: Multilingual Matters.

Wenden, A. L. (1985). Learner strategies. *TESOL Newsletter, 19*(5), 1-7.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly, 17*(2), 165-187.

Acknowledgment

I wish to acknowledge Elaine Tarone, Michael McCarthy, Roberta Abraham, and Rebecca Oxford for helpful feedback on earlier drafts of this article, and Sonya Yoder for assistance in editing the paper.

Author

ANDREW D. COHEN, Director, Institute of Linguistics and Asian and Slavic Languages & Literatures, University of Minnesota, 196 Klaeber Court, 320 16th Ave SE, Minneapolis, MN 55455. Specializations: language learning strategies, foreign language education.

Applied Language Learning
1996, Vol. 7, Nos. 1 & 2, pp. 25-45

Employing a Questionnaire to Assess the Use of Language Learning Strategies

Rebecca L. Oxford
University of Alabama

Questionnaires are among the most efficient and comprehensive ways to assess frequency of language learning strategy use. This article discusses the validity of the most widely employed strategy questionnaire, the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). Validity of the SILL rests on its link with language performance (course grades, standardized test scores, ratings of proficiency), as well as its relationship to learning styles. Reliability of the SILL is high across many cultural groups. Appropriate uses and limitations of questionnaires for strategy assessment are detailed, along with implications for research and instruction.

One of the most prevalent ways to assess the use of language learning strategies is to use a questionnaire (otherwise known as an *inventory* or a *summative rating scale*). The strategy questionnaire most often used around the world at this time is the *Strategy Inventory for Language Learning* (*SILL*, Oxford, 1986-1990). This article has five purposes: (1) to discuss strategy questionnaires other than the *SILL*, (2) to describe the *SILL's* purpose and nature, (3) to provide detailed psychometric results concerning the ESL/EFL (English as a second or foreign language) version of the *SILL*, (4) to present information on the appropriate uses and limitations of a strategy questionnaire in comparison with other means of strategy assessment, and (5) to provide implications for research and instruction.

Strategy Questionnaires Other than the SILL

To present a context, we turn first to strategy questionnaires other than the *SILL*. Nearly a dozen have been used in published studies. For example, Bialystok (1981) used a 12-item, structured, untitled rating scale to assess strategy use. The scale asked questions about the extent to which strategies were used on both oral and written tasks in communicative settings (the strategies were functional practice and inferencing or guessing) and in formal classroom settings (the strategies were formal practice and monitoring). Using the scale with students of French in grades 10 and 12 in Canada, Bialystok found that functional practice had a stronger relationship with achievement

© 1996, Rebecca L. Oxford

than did any of the other strategies, even though monitoring and inferencing were used more often. Formal practice with rules and structures was less effective as students advanced to higher levels of learning, but functional practice had no such limitation. Reliability and validity data were absent for this instrument.

Politzer (1983) published an untitled, 1-4-scaled strategy scale including 51 items divided into three groups: general behaviors, classroom behaviors, and interactions outside of class. Using this survey with U.S. university students of French, German, and Spanish, Politzer found that course level influenced strategy use, with higher-level students using more so-called "positive" strategies (i.e., strategies related to communicative language proficiency); and that females used social learning strategies more often than males. No reliability or validity data were given.

Politzer and McGroarty (1985) used a somewhat similar *Behavior Questionnaire* containing 66 items divided into three groups: individual study behaviors, classroom behaviors, and interactions outside of class. Reliability was marginally acceptable (.51, .61, and .63). The survey was used with students learning intensive ESL in an eight-week course. Improvements in ESL achievement were related to individual strategies, such as asking questions for clarification. Successful strategies for grammar differed from those for listening and speaking. Major academic field had a significant effect on strategy choice, with engineers avoiding strategies that were deemed "positive" for gaining communicative language proficiency; but there was an overlap with nationality, since many engineers were also Asian.

McGroarty (1987) used a 56-item *Language Learning Strategy Student Questionnaire* with a 0-6 range, divided into the same three groups as in the Politzer and McGroarty study above. No reliability or validity data were published. University students of Spanish, although taught by communicative methods, nevertheless avoided authentic practice strategies and used traditional learning strategies, such as relying heavily on the dictionary.

The *Learning Strategies Inventory* (Chamot, O'Malley, Kupper, & Impink-Hernandez, 1987) is a 48-item, 1-4-scaled instrument divided into five parts: listening in class, speaking in class, listening and speaking outside of class, writing, and reading. The items reflected a variety of ways of applying a total of 16 strategies. Results showed that students of Russian used more strategies than students of Spanish, while Spanish and Russian students used somewhat different strategies across language levels (beginning and intermediate or advanced). No data were published on reliability or validity.

Padron and Waxman (1988) developed a 14-item, 1-3-scaled instrument to assess reading strategies of Hispanic ESL students in grades 3-5. Seven of the items were expected to be positively related to learning and seven negatively related. Results showed that six of the seven most-used strategies were in the predicted-positive group. However, only two strategies were significantly related to learning outcomes, and these were both in the negative direction; no strategies significantly helped learning to occur. No reliability or validity data were offered.

Bedell (1993) points out a number of additional strategy scales. Huang (1984) and Huang and van Naerssen (1987) used a *Strategies Questionnaire* for Chinese EFL learners. This instrument includes some scaled items and some yes-no items, as well as free-response questions. Most of the items concern strategies for improving listening and speaking skills. Wangsotorn, Sripaipan, Rattanaprucks, Jarunggidanan, Singkalwanij, and Vejaphurti (1986) used the *Chulalongkorn University Language Institute Learning Strategy Form A* (consisting of 42 yes-no statements about students behaviors) for Thai learners of EFL. Kim (1991) designed a *Perceptual Learning Strategy Questionnaire*, including 18 items. Noguchi's (1991) *Questionnaire for Learners* is an instrument with 24 items on a 3-point scale followed by 24 on a 4-point scale, based largely on items from the *SILL*. Wen and Johnson's (1991) strategy scale is also adapted from the *SILL*.

Few of the above instruments have any published reliability or validity data. This is the key reason that the *SILL* was developed. If the psychometric properties of reliability and validity have not been explored, it is impossible to know whether we can put faith in the results of the research. Another reason for developing the *SILL* is that the preceding instruments do not always systematically represent the wide variety of strategies viewed as important to language learning; often they stop with cognitive and metacognitive strategies. Thus a more comprehensive scale was needed for measuring strategy use among ESL and EFL students.

The Strategy Inventory for Language Learning (SILL)

Development

The *SILL* (Oxford, 1986-1990) was first designed as an instrument for assessing the frequency of use of language learning strategies by students at the Defense Language Institute Foreign Language Center in Monterey, California. Two revised versions of the *SILL*—one for foreign-language learners whose native language is English (80 items) and the other for learners of English as a second or foreign language (ESL/EFL, 50 items)—were published in an appendix to Oxford's (1990b) learning strategy book for language teachers. This article deals only with *research done using the 50-item (short) version*. For details on the longer version, see Ehrman and Oxford (1989, 1990), Nyikos and Oxford (1993), Oxford (1986), Oxford and Ehrman (1995), Oxford and Nyikos (1989), Wildner-Bassett (1992a), and Bedell (1993).

It is estimated that 40 to 50 major studies, including a dozen dissertations and theses, have been done using the *SILL*. These studies have, by late 1995, involved approximately 10,000 language learners. According to research reports and articles published in the English language within the last ten to fifteen years, the *SILL* appears to be the only language learning strategy questionnaire that has been extensively checked for reliability and validated in multiple ways.

The *SILL* uses a choice of five Likert-scale responses for each strategy described: *never or almost never true of me*, *generally not true of me*, *somewhat true of me*, *generally true of me*, and *always or almost always true of me*. The *SILL* response

options are based on the widely used and well accepted response options of the *Learning and Study Strategies Inventory* described by Weinstein, Palmer, and Schulte (1987). On the *SILL*, learners are asked to indicate their response (1, 2, 3, 4, or 5) to a strategy description, such as "I try to find patterns in English" or "I plan my schedule so I will have enough time to study English." In addition to the original English version, the ESL/EFL *SILL* has been translated and used in the following languages: Arabic, Chinese, French, German, Japanese, Korean, Portuguese, Russian, Spanish, Thai, and Ukrainian.

Two strategy experts matched the *SILL* items with agreement at .99 against entries in a comprehensive strategy taxonomy of language learning. This taxonomy was built from a detailed blueprint of a range of over 200 possible strategy types (for complete details see Oxford, 1986).

One important note is that the *SILL* conceptualizes language-learning strategies in a broad way to include the social and affective sides of the learner as well as the more intellectual (cognitive) and "executive-managerial" (metacognitive). Therefore, when the *SILL* is related to language performance, the "whole learner," rather than just the cognitive and metacognitive aspects of the learner, is usually involved. This implies that language learning, as much as or more than almost any other discipline, is an adventure of the whole learner, not just a mental exercise.

In 1989, the *SILL* was organized according to strategy groups using a factor analysis. This procedure allowed the researcher to divide the instrument into dimensions usually referred to as subscales or factors. Six subscales were developed based on the early factor analyses, with the intent that each subscale would have an adequate number of items to facilitate more in-depth understanding of the learning strategies for ESL/EFL. These subscales included:

1. *Memory strategies*, such as grouping, imagery, rhyming, and structured reviewing (9 items).
2. *Cognitive strategies*, such as reasoning, analyzing, summarizing (all reflective of deep processing), as well as general practicing (14 items).
3. *Compensation strategies* (to compensate for limited knowledge), such as guessing meanings from the context in reading and listening and using synonyms and gestures to convey meaning when the precise expression is not known (6 items).
4. *Metacognitive strategies*, such as paying attention, consciously searching for practice opportunities, planning for language tasks, self-evaluating one's progress, and monitoring errors (9 items).
5. *Affective (emotional, motivation-related) strategies*, such as anxiety reduction, self-encouragement, and self-reward (6 items).
6. *Social strategies*, such as asking questions, cooperating with native speakers of the language, and becoming culturally aware (6 items).

As shown above, the largest group of items is the cognitive strategies. This stands to reason, because research on learning strategies suggests that cognitive strategies possess the greatest variety, covering strategies related to practice and to the all-important "deep processing" in which learners analyze, synthesize, and transform new information (Oxford & Ehrman, 1995).

A *SILL* package includes a short set of directions to the student with a sample item, the 50-item instrument, a scoring worksheet on which students record their answers and calculate their averages for each strategy subscale and their overall average, a summary profile that shows their results and provides examples for self-interpretation, and a strategy graph that allows each learner to graph results from the *SILL*. A background questionnaire is also available to document age, sex, language experience, motivation, and other information (see Oxford, 1990b).

Psychometric Qualities of the ESL/EFL SILL

This section describes the psychometric qualities of the 50-item ESL/EFL *SILL*. Normally, such quality is established and presented in terms of reliability and validity. (Note that psychometric quality data are also available for the longer form of the *SILL* that was designed for native English speakers learning foreign languages; see especially Oxford, 1992 and Oxford & Ehrman, 1995.)

Reliability

Reliability refers to the *degree of precision or accuracy of scores on an instrument*. In the case of the *SILL*, Cronbach alpha, a measure of internal consistency, was chosen as the most appropriate reliability index. The Cronbach alpha reliability coefficient is used on continuous data such as the Likert-type scale in the *SILL*.

Though the current ESL/EFL *SILL* was constructed using six subscales, reliability of the *SILL* is determined with the whole instrument. This is because the six subscales are strongly correlated with the *SILL* mean (.66 to .81) and moderately correlated with each other (.35 to .61); see Oxford and Ehrman (1995).

In general, the ESL/EFL *SILL* reliabilities have been high. With the ESL/EFL *SILL*, Cronbach alphas have been .94 using the Chinese translation with a sample of 590 Taiwanese university EFL learners (Yang, 1992a); .92 using the Japanese translation with 255 Japanese university and college EFL students (Watanabe, 1990); .91 using the Korean translation with 59 Korean university EFL learners (Oh, 1992); .93 using the researcher-revised Korean translation with 332 Korean university EFL learners (Park, 1994); and .91 using the Puerto Rican Spanish translation with 374 EFL learners on the island of Puerto Rico. (These reliabilities are similar to the range of .91 to .95 found for the 80-item foreign language *SILL* given in the native language of the respondent; see Bedell, 1993; Ehrman & Oxford, 1989, 1990; Nyikos & Oxford, 1993; Oxford, 1986; Oxford & Burry, 1993; Oxford & Ehrman, 1995; Oxford & Nyikos, 1989; Wildner-Bassett, 1992a).

Slightly lower but still very acceptable reliabilities are found for the ESL/EFL *SILL* when it is *not* administered in the native language of the respondents but is given in English instead. All the reliabilities in this paragraph refer to heterogeneous (multi-language) groups of ESL learners in the U.S. Phillips' (1990, 1991) data had a reliability of .87 with 141 students. *SILL* data from Oxford, Nyikos, Nyikos, Lezhnev, Eyring, and Rossi-Le (1989) showed a reliability of .86 with 159 students. Anderson's (1993) data on 95 students had a reliability of .91. Involving 31 learners, Talbott's (1993) data had a reliability of .85. A three-study combination (merging ESL data from Anderson [1993]; Talbott [1993]; and Oxford, Talbott, & Halleck [1990]) showed a reliability of .88 with 137 students.

Thus, reliability of the ESL/EFL *SILL* goes down, but not greatly, when the *SILL* is administered in the target language, English, rather than in the respondent's native language. These reliabilities are very respectable, and the *SILL* can be administered in the respondent's native language or a foreign or second language with confidence that measurement error is minimal.

Validity

Validity refers to *the degree to which an instrument measures what it purports to measure.* In the past, several bases have existed for justifying validity: content validity, criterion-related validity (predictive or concurrent), and construct validity. At the current time, these aspects of validity have been condensed to a single general validity (Chapelle, 1994; Messick, 1989).

Justifications of this broad-scale validity, according to Messick (1989), are comprised of construct validity, utility, value implications, social consequences, interpretation, and real-world action. Chapelle (1994, p. 161) asserts that "construct validity is central to all facets of validity inquiry, as most researchers have agreed for some time" and that "researchers should also consider justifications [for instrument validity] pertaining to test utility and the consequences of testing. . . . Researchers are obligated to use construct validity evidence as a basis for considering how their instruments impact the contexts in which they are used." This makes validation an ongoing process, which is the opposite of the notion of an instrument "validated once and for all time." Evidence of validation is not singular but additive. Further, it is not just the instrument that is validated, but also its use in a far bigger context of interpretation and action.

The data on the *SILL* show an ongoing effort at validation, using evidence based partially on construct validity.

Construct Validity as Seen in Relationships between Strategies and Language Performance

SILL construct validity is partially shown in relationships between the *SILL* on the one hand and language performance on the other. This evidence is probably the strongest support possible for the assertion of the construct validity of the *SILL*.

A number of ESL/EFL studies have demonstrated this relationship. In these

studies, language performance is measured in various ways: general language proficiency tests (Rossi-Le, 1989; Wen & Johnson, 1991; Green & Oxford, 1992; Phillips, 1990, 1991; Chang, 1991; Park, 1994), oral language proficiency tests (Chang, 1991), grades in a language course (Mullins, 1991), language achievement tests directly related to course content (Oxford, Park-Oh, Ito, & Sumrall, 1993a, 1993b), proficiency self-ratings (Oxford & Nyikos, 1989; Watanabe, 1990; Chang, 1991), and professional language career status (Ehrman & Oxford, 1989).

Here are some examples of the relationship between strategy use on the ESL/EFL *SILL* and language performance. Rossi-Le (1989) found that for 147 adult ESL students in the Midwestern and the Northeastern parts of the U.S., language proficiency level (on a standardized test) predicted strategy use in multiple-regression analyses. More proficient ESL students used self-management strategies like planning and evaluating ($p < .006$) and formal practice ($p < .02$) significantly more often than less proficient ESL students.

Strategy use was related to language achievement scores (final test grades) in a study involving 107 high school students of Japanese. The ESL/EFL *SILL* was modified slightly for the distance education students in this study by Oxford, Park-Oh, Ito, and Sumrall (1993a, 1993b). In a multiple regression analysis, learning strategy use was a moderate but significant predictor of Japanese language achievement (.20, $p < .04$). The only other significant predictor was the degree of learner motivation (.30, $p < .003$). (Note that this study adapted the ESL/EFL *SILL* for a group learning a language other than English.)

Using a modified Japanese version of the ESL/EFL *SILL*, Wen and Johnson (1991) studied the learning strategies of 242 second-year English majors at seven post-secondary institutions in Nanjing and Shanghai. These subjects had recent national English proficiency scores that averaged 10 points higher than the country's mean. Using partial least squares, the researchers found that one-third of the variance in English proficiency was related to combined effects of six variables, three of which were groups of strategies taken from the *SILL*.

Takeuchi (1993a) used multiple regression and found that eight *SILL* items predicted 58% of the variance in scores on the *Comprehensive English Language Test (CELT)*. The *CELT* was used in that study to measure English achievement among 78 Japanese first-year students of English at a women's college in Kyoto. The figure of 58% is unusually high for just eight strategies. Four strategies positively predicted language achievement: writing notes, messages, letters, or reports in English; trying not to translate verbatim; dividing words into parts to find meaning; and paying attention when someone is speaking English. This means that students who used English for multiple forms of writing, avoided translation, used word-analysis, and paid close attention were more likely to have high *CELT* scores. Four strategies negatively predicted language achievement: asking questions in English; using flashcards; writing down feelings in a language learning diary; and trying to find as many ways as possible to use English. In other words, the more these strategies were used, the lower the *CELT* score.

This is explainable partly because lower-proficiency students might ask more questions, have more turbulent feelings and be more willing to write them down, use very basic strategies like flashcards, or reach out (perhaps in desperation) to many different ways of using English. Takeuchi (1991a, 1991b, 1993a, 1993b) explained some of these findings based on cultural influences.

Watanabe (1990) asked university and college EFL students in Japan to rate from low to high their own proficiency in English. These proficiency self-ratings correlated moderately (average $r = .30$) with *SILL* strategies ($p < .0005$-$.001$), except for those in the category of social/affective strategies. This trend indicates that most *SILL* strategies were used more often by students who rated their language proficiency higher and they were used less often by students who rated their language proficiency lower.

Chang (1991) used the *SILL* to investigate the learning strategies and English proficiency of 50 mainland Chinese and Taiwanese ESL students at a southeastern university in the U.S. Three measures of proficiency (self-ratings and two standardized tests) showed different statistical effects on strategy use. Students who rated themselves above average in proficiency used more strategies overall than those who rated themselves below average. Neither the scores on the *Test of English as a Foreign Language* (*TOEFL*) nor the *Ilyin Oral Interview* were significantly related to overall strategy use, but students with high scores on the oral interview used significantly more social strategies than those with low scores.

Park (1994) employed the *SILL* to determine the relationship between strategy use and proficiency among 332 students of EFL at the Korea Maritime University and Inha University. Park divided the subjects into three groups according to their strategy use: low, medium, and high. Then Park calculated *TOEFL* scores for each group. According to an ANOVA, the *TOEFL* mean scores of these three groups differed significantly from each other. Post-hoc tests showed that the high strategy use group had a language proficiency score that was significantly higher than that of the medium strategy use group, which in turn had a slightly higher language proficiency score than that of the low strategy use group. Thus, a linear relationship was shown between strategy use and language proficiency. In addition, Park found that the correlation between total *TOEFL* scores and strategy use was $r = .34$ ($p < .0001$). Cognitive, social, and metacognitive strategies had a slightly higher relationship ($r = .33$, $.30$, and $.28$ respectively) to *TOEFL* scores than did other kinds of strategies (memory, $r = .24$; affective, $r = .23$; compensation, $r = .21$).

Phillips (1990, 1991) found strong relationships between ESL/EFL *SILL* frequencies and English proficiency levels (measured by the *TOEFL*) among 141 adult ESL learners in seven western states in the U.S. She found no consistent differences between high-proficiency students and low-proficiency students on entire strategy categories, so she looked at strategies singly. She found that middle scorers on the *TOEFL*, who thus had moderate proficiency in English, showed significantly higher overall strategy use than did the high-proficiency or the low-proficiency group, when strategy use was defined as the mean number of strategies used frequently and the mean number of strategy categories that had at least one frequently used strategy. The profile of

medium-proficiency students using more strategies more often than high-proficiency or low-proficiency students produced a curvilinear pattern. Additionally, Phillips discovered that high *TOEFL* scorers used such learning strategies as paraphrasing, defining clear goals for learning English, and avoiding verbatim translation significantly more often than low *TOEFL* scorers. The low *TOEFL* scorers, many of whom would logically be found among beginning students, reported significantly greater use of such strategies as using flashcards, finding out how to be a better speaker, looking for conversation partners, noticing tension or nervousness, and writing down feelings in a journal.

Green (1991) investigated 213 Spanish-speaking students learning English on the island of Puerto Rico. The *English as a Second Language Achievement Test* (*ESLAT*), which was used in the study, is a measure of overall English proficiency (not achievement on a given curriculum). Green found moderate and significant correlations, usually in the upper .30s, between *SILL* strategy factors and *ESLAT* proficiency scores, and he discovered the same level of correlations between individual *SILL* items and proficiency scores. In a later analysis of variance, Green (1992) showed that language level had a statistically significant influence on strategy use, with higher-proficiency students in general using strategies more frequently than lower-proficiency students. With a larger sample of 374 students, Green and Oxford (1995) found that language proficiency level had significant effects on the use of the following kinds of strategies: compensation strategies ($p < .0001$), cognitive strategies ($p < .0001$), metacognitive strategies ($p < .0025$), and social strategies ($p < .008$). Two other categories of strategies, memory and affective strategies, displayed no significant difference by proficiency level. In the four significant categories, higher proficiency was associated with more frequent strategy use. Significant variation occurred by gender, with females using strategies significantly more often than males in this study.

In Mullins' (1991) *SILL* study, 110 Thai university-level EFL majors showed linkages between strategy use and various measures of English proficiency. For instance, compensation strategy use correlated at $r = .38$ ($p < .0001$) with language placement scores and at $r = .32$ ($p < .006$) with language course grades. A correlation of $r = .24$ ($p < .03$) was found between metacognitive strategy use and language course grades. However, a negative correlation of $r = -.32$ ($p < .005$) was found between affective strategy use and language entrance examination scores, which are different from language placement scores in this particular Thai university. It is possible that students who are very anxious and who resort to affective strategies do less well on the entrance examination.

As shown by Dreyer and Oxford (1996), approximately 45% of the total variance in language proficiency (*TOEFL* scores) in a South African ESL study was explained by learning strategy use as measured by the *SILL*. A regression analysis demonstrated that the greatest part of the variance stemmed from metacognitive strategies, with much smaller amounts contributed by affective and social strategies. Canonical correlation showed a highly significant relationship between the parts of the *TOEFL* and the categories on the *SILL* ($r = .73$). The sample consisted of 305 Afrikaans first-year university students learning ESL in South Africa (Dreyer, 1992).

What we can learn about construct validity of the SILL based on relationships with language performance.

ESL/EFL *SILL* strategy use is related, as expected, to language performance in a number of studies, thus providing construct validity evidence for the *SILL*. (These results agree with earlier research using varied strategy-assessment instruments; for instance, Corrales & Call, 1989; Huang, 1984; Oxford & Nyikos, 1989; O'Malley & Chamot, 1990). In many, but not all instances, the relationship is linear, showing that more advanced or more proficient students use strategies more frequently.

Construct Validity in Strategy Use in Foreign versus Second Language Environments

Some existing *SILL* data indicate that strategy-use patterns often differ between ESL and EFL settings. ESL environments typically show high frequencies of use for at least half of the strategy categories. For example, Oxford, Nyikos, Nyikos, Lezhnev, Eyring, and Rossi-Le (1989) found high frequencies of use for 60% of the strategies on the *SILL* as used by 159 ESL learners in the U.S. Rossi-Le (1989) learned that among 147 adult ESL learners in two community colleges in the U.S., high frequencies existed for most of the strategies. Oxford, Talbott, and Halleck (1989), with a sample of 43 ESL learners at a large Northeastern university in the U.S., discovered high levels of strategy use for two-thirds of the strategy categories. Phillips (1990, 1991), studying 141 adult ESL learners in seven western states, noted that half the strategy categories were used at a high level. All four of these ESL studies reflected large amounts of high-frequency strategy use.

In contrast, five EFL studies (Klassen [1994], 228 Taiwanese university students; Noguchi [1991], 174 Japanese junior high students; Oh [1992], 59 Korean university students; Park [1994], 332 Korean university students; Yang [1994], 68 Taiwanese university students) indicated that strategy use in these settings was mostly at a medium level, far different from ESL strategy use in the studies noted earlier. In a study in Puerto Rico, where English was not the major language of daily communication but was nevertheless highly available as input (a hybrid ESL/EFL setting), strategy use among 213 university students learning English was much more like the EFL settings than the ESL environments.

Thus we can see that a *second* language environment, which demands daily use of the target language, often calls for (or encourages) more frequent strategy use than a *foreign* language environment, which does not require continual use of the target language. This is a sound generalization for most language students.

However, career interests can override this generalization. For example, Mullins (1992) found that her 110 Thai university EFL students had a rather high level of strategy use; and these students, unlike most of the EFL students mentioned in other studies, were majoring in English and wanted to use English in their careers. Likewise, Davis and Abas' (1991) 64 EFL students had mostly high strategy use; and they were already language teachers. Thus, strong language-related career interests can transform the target-language-impoverished EFL setting and make it a thriving home for language

learning strategies.

What we can learn about construct validity of the SILL based on relationships with ESL/EFL setting.

In short, unless foreign language students are extremely motivated because of their language-career interests, they will use strategies with less frequency than second-language students. Second-language students have more daily need to use the language, and therefore it is likely that they have greater motivation to use strategies. (Strategy use and motivation have elsewhere been shown to be statistically related by Oxford & Nyikos, 1989.)

Construct Validity in the Relationship Between Strategies and Learning Styles

Strong relationships between learning strategy use and sensory preferences—often viewed as an aspect of learning style—have been posited (Oxford, Ehrman, & Lavine, 1991) as partial evidence of the construct validity of the *SILL*. According to Oxford, Ehrman, and Lavine, visual students use strategies involving reading alone in a quiet place or paying attention to blackboards, movies, computer screens, and other forms of visual stimulation. Auditory students are comfortable without visual input and frequently use strategies that encourage conversation in a noisy, social environment with multiple sources of aural stimulation. Kinesthetic students need movement strategies, and tactile students require strategies that involve manipulating real objects in the classroom; both types need to use the strategy of taking frequent breaks.

ESL/EFL *SILL* data exist supporting the link between learning strategy use and learning styles, thus at the same time strengthening the evidence of construct validity of the *SILL*. Rossi-Le (1989) found a significant relationship ($p < .0005$) between learning styles (visual, auditory, tactile, and kinesthetic) and overall strategy use on the ESL/EFL *SILL* through a MANOVA, and she also found significant predictive relationships through multiple regression.

Rossi-Le's MANOVA results showed that visual learners preferred visualization strategies ($p < .0005$). Auditory-style learners used memory strategies more than did other learners ($p < .0005$). Compared with others, tactile learners demonstrated significant use of strategies for searching for and communicating meaning ($p < .006$) and self-management/metacognitive strategies ($p < .02$). Kinesthetic learners did not use general study strategies ($p < .003$) or self-management/metacognitive strategies ($p < .02$) as often as others did.

The regression results indicated that a visual learning style predicted using visualization strategies (beta =.33, $p < .00005$). Being a visual learner, however, negatively predicted using independent strategies (beta = -.22, $p < .001$), affective strategies (beta = -.23, $p < .009$), and strategies for searching for and communicating meaning (beta =-.22, $p < .008$). Having an auditory learning style significantly predicted memory strategies (beta = .38, $p < .0008$) and self-management or metacognitive strategies (beta = .20, $p < .01$) but was a negative predictor of employing authentic language-use strategies (beta = -.20, $p < .01$). Being a tactile learner significantly predicted employing authentic language-use strategies (beta = .26, $p < .001$) and strategies for meaning (beta =.32,

.0002) but negatively predicted use of memory strategies (beta = -.16, p < .04). A kinesthetic learning style predicted infrequent use of general study strategies (beta = -.32, p < .002). Thus, these predictions are low-to-moderate and significant.

What we can learn about construct validity of the SILL based on relationships with learning styles.

Though existing evidence is sparse, the data we have indicate that learning strategy use is related to learning styles. It is as though learning styles are the underlying or internal construct, and learning strategies are the more "outward" manifestation of learning styles.

The relationship is by no means simple, however. Predictions of strategy use according to learning style are sometimes straightforward (e.g., visual learning style predicts visualization strategy use) and sometimes not so straightforward (e.g., auditory learning style predicts metacognitive strategy use). Clearly more information is needed on the links between learning styles and learning strategies.

Construct Validity in the Relationship Between Strategy Use and Gender

In many ESL/EFLstrategy studies, results have usually favored females as more frequent users of strategies (for instance, Dreyer, 1992; Ehrman & Oxford, 1989; Green, 1991, 1992; Green & Oxford, 1993, 1995; Noguchi, 1991; Oxford, 1993a, 1993b; Oxford, Ehrman, & Nyikos, 1988; Oxford, Park-Oh, Ito, & Sumrall, 1993a, 1993b; Yang, 1992b, 1993). In a few studies, females have had a distinctly different pattern of strategy use from that of males (Bedell, 1993; Watanabe, 1990). Some studies, noted by Bedell and by Green and Oxford, have shown that males surpassed females on a certain number of separate strategies but not on whole clusters or groups of strategies.

What we can learn about construct validity of the SILL based on relationships with gender.

Overall, the last decade of studies has shown that females are generally more frequent strategy users than men in a language learning situation. This trend fits in with previous theory and research about females as better, more efficient learners and users of language (native or other) than males; see Oxford (1993a, 1993b) for many biosocial reasons for this difference. Thus, the construct validity of the *SILL* has additional evidence.

Other Aspects of Validity: Utility, Value Implications, Social Consequences, Interpretation, and Real-World Action

According to Messick (1989) and Chapelle (1994), aspects of general validity (in addition to construct validity) include utility, value implications, social consequences, interpretation, and real-world action.

Utility can be defined as the *usefulness of an instrument in real-world settings for making decisions relevant to people's lives.* The *SILL* has utility, deemed to be a crucial piece of evidence of general validity of the instrument. Utility is demonstrated by the many people around the world who have employed the *SILL* and by the uses to which they have put it. The most frequent venue of use has been the classroom, where

the goal has been chiefly to reveal the relationship between strategy use and language performance. This goal is important because if there is a strong relationship between these two variables, perhaps language performance can be improved by enhancing strategy use.

Related to utility are the value implications and social consequences of the questionnaire. Underlying the *SILL* is the value of learning strategies as tools for learner self-direction, autonomy, and achievement. Strategies are a means of enhancing learning for each student. Every student uses strategies, but some strategies are more appropriate than others to a given task and to the student's own learning style (visual, auditory, hands-on; extroverted, introverted; and so on). The social consequences of using the *SILL* are that learners (and their teachers) become more aware of the strategies each learner typically uses. This awareness helps teachers more effectively design language instruction and enables them to provide relevant strategy instruction. This awareness also helps students seek and experiment with new and more efficient strategies.

Interpretation and real-world action relate to outcomes of testing. Interpretation of the *SILL* should be limited to "typical" strategies of a given student (or, when aggregated, strategies of a given group) in a variety of situations and tasks. It should not be applied to assess the strategies used for a single activity, such as the very task-bound strategies that Marilyn uses to read aloud in French the first part of Chapter 3 of *Madame Bovary*.

Real-world actions based on the appropriate interpretation of the *SILL* include increased theory-building concerning the nature of language learning strategies; assessing strategy use at a given point, to be compared with strategy use later (sometimes after strategy improvement interventions); comparing strategy use with proficiency or achievement; comparing the learning strategies of women and men; making the conceptual linkage between strategy use and learning styles; and individualizing classroom instruction based on the strategy use of different students. So far the utility of the *SILL* has not included making placements of individuals into language classes on the basis of strategy use results, although such strategy information could conceivably be combined with other kinds of data for making such placements. See the reference list for dozens of studies showing various applications of the *SILL*.

Appropriate Uses and Limitations of Questionnaires Compared with Other Strategy Instruments

Compared with the other strategy assessment techniques (see Table 1), *student-completed strategy questionnaires* have a very important and appropriate use. These questionnaires provide a general assessment of each student's typical strategies across a variety of possible tasks. However, strategy questionnaires do not describe in detail the language learning strategies a student uses in response to any *specific* language task (as do some specific-strategy interviews or think-aloud protocols). For a researcher or a teacher who wants to discover strategy use on a particular reading

TABLE 1
Comparisons of strategy-assessment types

Type of assessment	Appropriate uses	Limitations of use
Strategy questionnaires	Identify "typical" strategies used by an individual; can be aggregated into group results; wide array of strategies can be measured by questionnaires	Not useful for identifying specific strategies on a given language task at a given time
Observations	Identify strategies that are readily observable for specific tasks	Not useful for unobservable strategies (e.g., reasoning, analyzing, mental self-talk) or for identifying "typical" strategies
Interviews	Identify strategies used on specific tasks over a given time period or more "typically" used strategies; usually more oriented toward task-specific rather than "typical" strategies of an individual; depends on how interview questions are asked	Usually less useful for identifying "typical" strategies because of how interviews are conducted, but could be used for either task-specific or "typical" strategies
Dialogue journals, diaries	Identify strategies used on specific tasks over a given time period	Less useful for identifying "typical" strategies used more generally
Recollective narratives (language learning histories)	Identify "typical" strategies used in specific settings in the past	Not intended for current strategies; depends on memory of learner
Think-aloud protocols	Identify in-depth the strategies used in a given, ongoing task	Not useful for identifying "typical" strategies used more generally
Strategy checklists	Identify strategies used on a just-completed task	Not useful for identifying "typical" strategies used more generally

comprehension task in a given classroom on Monday morning, a general strategy questionnaire like the *SILL* would not be useful. It is a misuse of the *SILL* (or any other strategy questionnaire) to try to identify task-specific strategies with that instrument.

Strategy questionnaires have certain advantages. They are quick and easy to administer, may be the most cost-effective mode of strategy assessment, and are almost completely nonthreatening when administered using paper and pencil (or computer) under conditions of confidentiality. Moreover, many students discover a great deal about themselves from taking a strategy questionnaire, especially one like the *SILL* that is self-scoring and that provides immediate learner feedback.

An advantage specifically accruing to the *SILL* is that this questionnaire is free of social desirability response bias. Social desirability response bias, or the tendency to answer in a way that the researcher "wants" one to answer, is usually identified by a moderate to high correlation between the *Marlowe-Crown Social Desirability Scale* and a given instrument like the *SILL*. No such correlation appeared in a large-scale study by Yang (1992b), in which the researcher tested 505 students of EFL on the *SILL* and the *Marlowe-Crown*. Therefore, students appeared to express themselves freely and openly on the *SILL*. In other studies, the current author compared results of informal strategy interviews with the way that respondents answered on the *SILL* and found that respondents had answered the *SILL* honestly (Oxford, 1986). At this writing, no other language learning strategy questionnaire has been studied for social desirability response bias.

Implications for Research and Instruction

First, language researchers must conceptualize language learning strategies in a way that includes the social and affective sides of learning (as shown in the *SILL*) as well as the more intellectual and "executive-managerial" sides. Language learning is not just cognitive and metacognitive. It involves much more from the learner.

Second, through strategy assessment teachers can help their students recognize the power of using language learning strategies for making learning quicker, easier, and more effective. Teachers need to know the appropriate uses and limitations of each assessment technique, as seen in the previous section. Multiple techniques are to be encouraged whenever the time and resources are available. When time and resources are restricted, teachers should use the most reliable and valid strategy assessment measure that they can for the purposes they have defined. When the purposes include tapping the "typical" or general strategy use of an individual student or a group, strategy questionnaires like the *SILL* can be extremely helpful. If much more precise measurement of highly task-based strategy use is the purpose, then other measurement tools are required.

Third, based on the information from strategy assessment, teachers can weave strategy instruction into regular classroom events in a natural, comfortable, but explicit way. Chamot and Kupper (1989), Oxford (1990b), and O'Malley and Chamot (1990) provide helpful details on how to do this. Teachers must also keep in mind differences in motivation, learning style, gender, and other factors that affect learning strategy use.

Fourth, teachers need to be judicious in their selection of strategies to use in instruction, and existing research can provide good clues for this selection. Research indicates that some strategies in certain studies do *not* relate strongly to proficiency.

For example, based on Takeuchi's research (1993a), it is possible to say that using flash cards is clearly *not* a sure-fire strategy to promote proficiency in all cultures and for all kinds of learning styles. Flashcards might work for some students but not for others. On the other hand, research shows that paying attention and actively using the language for writing seem to be widely appropriate strategies in most contexts and for most kinds of learners.

Fifth, strategy assessments using different measurement modes with the same sample of students could be cross-correlated. This would contribute to the validity of various assessment techniques. For instance, it would be useful to correlate results from a think-aloud protocol, an interview, and a survey to see how closely they relate to each other. If results show that an interview and a survey are highly correlated but that they are only weakly correlated with a think-aloud procedure, this information would be useful in selection of an assessment procedure next time.

Sixth, studies will need to be replicated so that more consistent information becomes available within and across populations. Particularly important is more information on how students from different cultural backgrounds and different countries use language learning strategies. Teachers need to have more background on how to use such information in the classroom. Here is a clear opportunity for researchers to better translate their findings into materials to be used in the classroom.

In sum, it is critical that learning strategies be considered when planning courses, teaching students, and designing classroom research. Appropriate learning strategies should be among the first considerations of any ESL/EFL teacher or researcher who wants to enhance student learning.

References

Anderson, N. (1993). Data on adult ESL strategy use. Unpublished manuscript, Ohio University, Athens, OH.

Bedell, D. (1993a). Chinese data on EFL strategy use among university students. Unpublished manuscript, University of Alabama, Tuscaloosa, AL.

Bedell, D. (1993b). Crosscultural variation in the choice of language learning strategies: A mainland Chinese investigation with comparison to previous studies. Unpublished master's thesis, University of Alabama, Tuscaloosa, AL.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*, 69-83.

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *Modern Language Journal, 65*, 24-35.

Boraie, D., Kassabgy, O., & Oxford, R. (1994). Empowering teachers and learners: Style and strategy awareness. Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, Baltimore, MD.

Chamot, A. U., & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals, 22*(1), 13-24.

Chamot, A. U., O'Malley, J. M., Kupper, L., & Impink-Hernandez, M. (1987). *A study of*

learning strategies in foreign language instruction: First year report. InterAmerica Research Associates, Rosslyn, VA.

Chang, S-J. (1991). A study of language learning behaviors of Chinese students at the University of Georgia and the relation of those behaviors to oral proficiency and other factors. Unpublished doctoral dissertation, University of Georgia, Athens, GA.

Chapelle, C. (1994). Are C-tests valid measures for L2 strategy research? *Second Language Research, 10* (2), 157-187.

Cohen, A. D. (1987). Studying learner strategies: How we get the information. In A. L. Wenden, & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 31-42). Englewood Cliffs, NJ: Prentice-Hall.

Corrales, O., & Call, M.E. (1989). At a loss for words: The use of communication strategies to convey lexical meaning. *Foreign Language Annals, 22*, 227-240.

Davis, E., & Abas, H. (1991). *Second language learning strategies utilized by some members of language departments at four institutions - Sulawesi, Indonesia.* Sulawesi, Indonesia: Summer Institute of Linguistics.

Dreyer, C. (1992). Learner variables as predictors of ESL proficiency. Unpublished doctoral dissertation, Potchefstroom University, South Africa.

Dreyer, C., & Oxford, R. (1996). Learner variables related to ESL proficiency among Afrikaans-speakers in South Africa. In R. L. Oxford (Ed.), *Language learning strategies around the world.* Manoa: University of Hawai'i Press.

Ehrman, M. E., & Oxford, R. L. (1989). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *Modern Language Journal, 73*, 1-13.

Ehrman, M. E., & Oxford, R. L. (1990). Adult language learning styles and strategies in an intensive training setting. *Modern Language Journal, 74*(3), 311-327.

Green, J. M. (1991). Language learning strategies of Puerto Rican university students. Paper presented at the annual meeting of Puerto Rico Teachers of English to Speakers of Other Languages, San Juan, PR.

Green, J. M. (1992). Additional analyses of Puerto Rican strategy data. Unpublished manuscript, University of Puerto Rico at Mayaguez.

Green, J., & Oxford, R. (1993). New analyses on expanded Puerto Rican strategy data. Unpublished manuscript, University of Puerto Rico at Mayaguez and University of Alabama, Tuscaloosa, AL.

Green, J., & Oxford, R. (1995). A closer look at language learning strategies, L2 proficiency, and gender. *TESOL Quarterly, 29*, 261-297.

Huang, X. H. (1984). An investigation of learning strategies in oral communication that Chinese EFL learners in China employ. Unpublished master's thesis, Chinese University of Hong Kong.

Huang, X. H., & Van Naerssen, M. (1987). Learning strategies for oral communication. *Applied Linguistics, 8*, 287-307.

Keefe, J. W., & Monk, J. S., with Letteri, C. A., Languis, M., & Dunn, R. (1989). *Learning style profile.* Reston, VA: National Association of Secondary School Principals.

Kim, J-D. (1991). A comparison of learning strategies of college students enrolled in

beginning and advanced English as a second language. Unpublished doctoral dissertation, Loma Linda University, Loma Linda, CA.

Klassen, J. (1994). The language learning strategies of freshman English students in Taiwan: A case study. Master's thesis, California State University at Chico, CA.

Lawrence, G. (1984). A synthesis of learning style research involving the *MBTI*. *Journal of Psychological Type*, 2-15.

LoCastro, V. (1994). Learning strategies and learning environments. *TESOL Quarterly, 28*, 409-414.

McGroarty, M. (1987). Patterns of persistent second language learners: Elementary Spanish. Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, Miami, FL.

Messick, S. (1989). Validity. In R. E. Linn (Ed.), *Educational measurement* (3rd ed.). New York: Macmillan.

Mullins, P. (1992). Successful English language learning strategies of students enrolled in the Faculty of Arts, Chulalongkorn University, Bangkok, Thailand. Unpublished doctoral dissertation, United States International University, San Diego, CA.

Myers, I., & McCaulley, M. (1985). *A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Noguchi, T. (1991a). Review of language learning strategy research and its implications. Unpublished bachelor's thesis, Tottori University, Tottori, Japan.

Noguchi, T. (1991b). *Questionnaire for learners*. Tottori University, Tottori, Japan.

Nyikos, M., & Oxford, R. (1993). A factor analytic study of language-learning strategy use: Interpretations from information-processing and social psychology. *Modern Language Journal, 77*, 11-22.

Oh, J. (1992). Learning strategies used by university EFL students in Korea. *Language Teaching, 1*, 3-53.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Oxford, R. L. (1986). *Development and psychometric testing of the Strategy Inventory for Language Learning (SILL)*. ARI Technical Report 728. Alexandria, VA: Army Research Institute. Appendices as ARI Research Note 86-92.

Oxford, R. L. (1986-1990). *Strategy Inventory for Language Learning*. Various versions. Tuscaloosa, AL: Oxford Associates.

Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implications for strategy training. *System, 17*, 235-247.

Oxford, R. L. (1990a). Language learning strategies and beyond: A look at strategies in the context of styles. In S. S. Magnan (ed.), *Shifting the instructional focus to the learner* (pp. 35-55). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Oxford, R. L. (1990b). *Language learning strategies: What every teacher should know*. New York: Newbury House / Harper & Row. *Now* Boston: Heinle & Heinle.

Oxford, R. L. (1993a). Instructional implications of gender differences in language learning styles and strategies. *Applied Language Learning, 4*(1-2), 65-94.

Oxford, R. L. (1993b). *La différence continue . . . :* Gender differences in second/foreign language learning styles and strategies. In J. Sutherland (Ed.), *Exploring gender* (pp. 140-147). Englewood Cliffs, NJ: Prentice-Hall.

Oxford, R. L., & Burry, J. A. (1993). Evolution, norming, and psychometric testing of Oxford's *Strategy Inventory for Language Learning (SILL).* Paper presented at the annual meeting of the National Council on Measurement in Education, Atlanta, GA.

Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the *Strategy Inventory for Language Learning. System, 23*(1), 1-23.

Oxford, R. L., & Ehrman, M. E. (1989). Psychological type and adult language learning strategies: A pilot study. *Journal of Psychological Type, 16*, 22-32.

Oxford, R. L., & Ehrman, M. E. (1995). Adults' language learning strategies in an intensive foreign language program in the United States. *System, 23*(3), 359-386.

Oxford, R. L., Ehrman, M. E., & Lavine, R. Z. (1991). Style wars: Teacher-student style conflicts in the language classroom. In S. S. Magnan (Ed.), *Challenges in the 1990s for college foreign language programs* (pp. 1-25). Boston: Heinle & Heinle.

Oxford, R. L., & Green, J. M. (1995). Making sense of learning strategy assessment: Toward a higher standard of research accuracy—a reply to LoCastro's article. *TESOL Quarterly, 29*, 166-171.

Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students *Modern Language Journal, 73*(2), 291-300.

Oxford, R. L., Nyikos, M., & Ehrman, M. (1988). *Vive la différence?* Reflections on sex differences in use of language learning strategies. *Foreign Language Annals, 21*(4), 321-329.

Oxford, R. L., Nyikos, M., Nyikos, K., Lezhnev, V., Eyring, J., & Rossi-Le, L. (1989). Learning strategies of adult ESL learners in the U.S.: A collaborative study. Unpublished manuscript, University of Alabama, Tuscaloosa, AL.

Oxford, R. L., Park-Oh, Y., Ito, S., & Sumrall, M. (1993a). Factors affecting achievement in a satellite-delivered Japanese language program. *American Journal of Distance Education, 7*, 10-25.

Oxford, R. L., Park-Oh, Y., Ito, S., & Sumrall, M. (1993b). Learning a language by satellite: What influences achievement? *System, 21*(1), 31-48.

Oxford, R. L., Talbott, V., & Halleck, G. (1990). Language learning strategies, attitudes, motivation, and self-image of students in a university intensive ESL program. Paper presented at the annual meeting of International Teachers of English to Speakers of Other Languages, San Francisco, CA.

Padron, Y. N., & Waxman, H. C. (1988). The effects of ESL students' perceptions of their cognitive strategies on reading achievement. *TESOL Quarterly, 22*, 146-150.

Park, G. (1994). Language learning strategies: Why do adults need them? Unpublished manuscript, University of Texas at Austin.

Phillips, V. (1990). English as a second language learner strategies of adult Asian students using the *Strategy Inventory for Language Learning*. Unpublished doctoral dissertation, University of San Francisco, San Francisco, CA.

Phillips, V. (1991, Nov.). A look at learner strategy use and ESL proficiency. *CATESOL Journal,* 57-67.

Politzer, R. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievement. *Studies in Second Language Acquisition, 6*, 54-65.

Politzer, R., & McGroarty, M. (1985). An exploratory study of learning behaviors and their relationship to gains in linguistic and communicative competence. *TESOL Quarterly, 19*, 103-124.

Rossi-Le, L. (1989). Perceptual learning style preferences and their relationship to language learning strategies in adult students of English as a second language. Unpublished doctoral dissertation, Drake University, Des Moines, IA.

Talbott, V. (1993). Strategies of adult ESL students: Data. Unpublished manuscript, Skagit Valley Junior College, Washington.

Takeuchi, O. (1991a). A study of EFL students' use of language learning strategies. *Gohokenkyu to Eigokoyoiku (Studies on English Language Usage and Language Teaching), 13*, 58-65.

Takeuchi, O. (1991b). Language learning strategies in second and foreign language of strategies frequency. *Bulletin of the Institute for Interdisciplinary Studies of Culture, 8*, 64-83.

Takeuchi, O. (1993a). A study of language learning strategies and their relation to achievement in EFL listening comprehension. *Bulletin of the Institute for Interdisciplinary Studies of Culture, 10*, 131-141.

Takeuchi, O. (1993b). Language learning strategies and their relationship to achievement in English as a foreign language. *Language Laboratory, 30*, 17-34.

Wangsotorn, A., Sripaipan, T., Rattanaprucks, N., Jarunggidanan, T., Singkalwanij, P., & Vejaphurti, A. (1986). Relationships between learning modes and the beginners' success in English. Paper presented at the annual meeting of International Teachers of English to Speakers of Other Languages, Anaheim, CA.

Watanabe, Y. (1990). External variables affecting language learning strategies of Japanese EFL learners: Effects of entrance examination, years spent at college/university, and staying overseas. Unpublished master's thesis, Lancaster University, Lancaster, U.K.

Weinstein, C. E., Palmer, D., & Schulte, A. C. (1987). *Learning and Study Strategies Inventory.* Clearwater, FL: H & H Publishers.

Wen, Q., & Johnson, R. K. (1991). Language learning approaches and outcomes: A study of tertiary English majors in China. Paper presented at the Sixth International Conference of the Institute of Language in Education, Hong Kong, China.

Wenden, A., & Rubin, J. (Eds.). (1987). *Learner strategies in language learning.* Englewood Cliffs, NJ: Prentice-Hall.

Wildner-Bassett, M. (1992a). Data set of foreign language students' strategies. Unpublished manuscript, University of Arizona, Tucson, AZ.

Wildner-Bassett, M. (1992b). Relationships among language learning strategies, personality types, and learning styles. Paper presented at the annual meeting of the American Council on the Teaching of Foreign Languages, Chicago.

Yang, N-D. (1992a). Data on students in an intensive EFL program in Taipei. Unpublished manuscript, National Taiwan University.

Yang, N-D. (1992b). Second language learners' beliefs about language learning and their use of learning strategies: A study of college students of English in Taiwan. Unpublished doctoral dissertation, University of Texas, Austin, TX.

Yang, N-D. (1993). Understanding Chinese students' language beliefs and learning strategy use. Paper presented at the annual meeting of International Teachers of English to Speakers of Other Languages, Atlanta, GA.

Yang, N-D. (1994). An investigation of Taiwanese college students' use of English learning strategies. Research report, National Taiwan University, Taiwan.

Zhang, W. (1994). Data on university EFL language use. Unpublished manuscript, Ohio University, Athens, OH.

Author

REBECCA OXFORD, Associate Dean, College of Education, University of Alabama, 201 Carmichael Hall, Tuscaloosa, AL 35487. Specializations: language learning strategies, styles, motivation, and anxiety; psychometrics.

Applied Language Learning
1996, Vol. 7, Nos. 1 & 2, pp. 47-60

Validity Issues in Computer-Assisted Strategy Assessment for Language Learners

Carol A. Chapelle
Iowa State University

This article probes validity issues of computer-assisted strategy assessment (CASA) in second-language (L2) acquisition research. Using two CASA studies, it reviews the metacognitive and cognitive strategies. It discusses trait-oriented and interactionist approaches to the concept of strategy, and examines validity by focusing on "advance preparation" and "resourcing," examples of trait-oriented and interactionist strategies, respectively. It suggests additional forms of validity evidence and describes applications of CASA as well as problems associated with this form of measurement.

Researchers of L2 acquisition and computer-assisted language learning (CALL) have exploited computer technology as means of gathering performance data of learners working on language activities. These data are then used to make inferences about learners' linguistic competence (e.g., Bland, Noblitt, Armington, & Gay, 1990) and strategies for L2 acquisition. As computer technology becomes more widespread throughout educational settings, these research practices are likely to expand in their popularity. In anticipation that more researchers will be interested in CASA for L2 learners in the future, this article draws upon past experience to summarize pertinent issues.

CASA Research

Throughout the 1980's and early 1990's, a number of researchers have used computer-assisted methods for strategy research by collecting data generated by learners while using interactive software. In such studies, the recorded data were generated through the learner input to the computer program and the program responses to the learner. This work has not been reexamined and interpreted from a measurement perspective to identify common issues across individual studies and to define "CASA" as a method.

Although there are many possible roles for the computer in the learning environment and therefore in strategy research (e.g., Chun, 1994), the basic learner-computer interaction that occurs during individualized work has formed the basis for most L2 strategy research to date. I will focus on assessment of strategies familiar from other

© 1996, Carol A. Chapelle

L2 research rather than the fundamental cognitive processes that have been the object of inquiry in psychology laboratories for many years.[1]

When a computer program is used to control or facilitate a language task, it can unobtrusively document learners' behavior such as the time they spend at various points in a problem sequence, the order in which they complete steps, and the editing they do to produce a linguistic product. Such a program can collect as sequenced[2] data everything the learner sees on the screen, all keypresses and mouse-clicks that the learner makes, and the time that each action takes place. For the most part, software in such research has been designed or modified to capture and store the desired information. The question for the researcher is which of the obtainable data can be used as indicators of learner capacities. The CASA studies listed in Table 1 provide some tentative answers to this question.

One study of English as a second language (ESL) learners (Jamieson & Chapelle, 1987) investigated "advance preparation," a metacognitive strategy defined as "planning for and rehearsing linguistic components necessary to carry out an upcoming language task" (O'Malley et al., 1985, p. 33). Using ESL dictation tasks over the course of six weeks, learners listened to words (on the "spelling" task) and individual sentences (on the "dictation" task) and then typed what they had heard. In this setting, it was assumed that the time the learner spent after hearing the input and before responding was spent planning performance; therefore, advance preparation was inferred from the amount of time that elapsed between the end of the input signal and the time that the learner pressed the first key to begin to answer. The actual indicator of the degree to which each learner used advance preparation was obtained by having the computer store the time it took to respond to each item, and calculating the mean "time-to-begin" by dividing each learner's total by the total number of items she or he had completed. It is important to note that the learners in this study were not instructed to complete the exercise as quickly as possible; they were working routinely over the course of several weeks.

Another metacognitive strategy, monitoring output, was assessed in the same instructional setting. On the dictation tasks, learners were able to edit the response they had typed by deleting, inserting, and changing characters or words before the response was evaluated by the computer. This behavior was documented in the computer records; it was considered an indicator of monitoring output, in the sense described by Krashen (1982), Bialystok (1981), Wenden (1985), and O'Malley et al. (1985). The number of times a learner edited was divided by the total number of completed items to obtain the average number of times a learner edited each item.

A third metacognitive strategy investigated with the same software was "monitoring input," defined by Bialystok (1981) as, reflecting on the formal aspects of a message as it was comprehended. The dictation tasks allowed learners to listen to the input as many times as they wanted before typing it. The demands of the task required learners to focus on formal aspects of the input. When they had not comprehended a sentence or word the first time it was presented, they requested to hear it again

TABLE 1
Studies using computer-assisted strategy assessment

Assessment context		**Learner**	
Task goal	Behavior	Learner's assessment of task*	Strategy
construct & edit sentences [1]	completing extra sentences and experimenting with responses	instruction	exploration
construct & edit sentences [2]	requesting any help	instruction	resourcing
read & answer questions or summarize[3]	requesting dictionary	experiment definitions	resourcing*
complete grammar exercise [4]	requesting rule/example help	instruction	inferencing/ hypothesis-testing
complete grammar exercise [4]	requesting/not requesting help	instruction	controlled/ automatic processing
dictation: listen & write [5] word/sentence	latency before responding	instruction	advance preparation
dictation: listen & write [5] word/sentence	editing response	instruction	monitoring output
dictation: listen & write [5] word/sentence	requesting repeated input	instruction	monitoring input

[1] Hsu, Chapelle, & Thompson (1993)
[2] Chapelle & Mizuno (1989)
[3] Hulstijn (1993)
[4] Doughty (1987)
[5] Jamieson & Chapelle (1987)

* This refers to the learner's perception of what he/she is doing while participating in the activity. The learner's perception is believed to affect performance, and therefore should be considered in the interpretation of the performance. For example, "instruction" means the learner perceived the task as taking place for instructional purposes.

one or more times. Subsequently, those requests were used as evidence for monitoring input. In other words, "monitoring input" was inferred on the basis of observed requests.

Other interaction observation programs have been used to measure "resourcing" defined by O'Malley, et al. (1985) as the cognitive strategy of using reference materials to obtain information about the L2. Consistent with this definition, researchers have considered each request for on-line help as evidence for the use of resourcing. In one study (Hulstijn, 1993), learners had access to an on-line dictionary that they could consult while reading a passage followed by questions. The resulting data documented the words that they had looked up in the dictionary and the time that each was looked up. Also while investigating resourcing, Chapelle and Mizuno (1989) collected data as learners worked on a task requiring them to construct and edit sentences in the exercises. As learners worked, help was given only at their request; they could ask for help with vocabulary, grammar, or the semantic facts pertaining to the sentences.

Doughty and Fought (1984) also documented learners' help requests on grammar items but interpreted them differently. Learners' requests for help while they were used as indicators of learners' "controlled access of explicitly learned knowledge." In addition, the researchers reasoned that "attempts to complete tasks without any help from the program [they reasoned] reflect automatic access to implicit knowledge in memory" (Doughty, 1987, p. 151). Other strategies evidenced by learners using particular types of help in software were "hypothesis-testing" and "inferencing based on L1." Doughty and Fought operationalized the definitions of these strategies as the type of grammar help requested by learners. When learners consulted help consisting of examples, they were considered to be displaying evidence of "hypothesis-testing." When they chose grammar help consisting of formal rules or the correct answer, learners were considered to be "inferencing [about syntactic forms in the L2] based on L1."

Using a sentence-constructing and -editing task, Hsu, Chapelle and Thompson (1993) assessed another strategy—exploration, the use of software to experiment and test hypotheses about the target language (Higgins & Johns, 1984). Exploration was operationalized in two ways: the number of sentences learners constructed after having completed the number required by their assignment and the number of times they edited an answer after receiving a message that it was correct—which the software allowed but did not require them to do. The operational definition of exploration in this study was derived from the theoretical definition that included the notion that learners would be motivated and interested in experimenting with the software.

In summary, a number of strategies have been investigated using computer-assisted methods. To understand CASA as a method, however, it is informative to interpret the types of strategies from a measurement perspective and, consequently, to explore the nature of the strategy construct.

Definition of "Strategy" in Computer-Assisted Strategy Research

To examine this method of assessment, I consider strategies to be theoretical constructs that are themselves not observable, but are hypothesized to be responsible for observed behavioral data.[2] Since there is more than one way to define a theoretical construct, I examine carefully how a particular strategy such as "monitoring" is defined in a research study. By looking for similarities in strategy definition across computer-assisted research studies, I query the extent to which an approach to strategy definition is tied to CASA. I provide means for considering the "nature of the construct," by distinguishing two approaches to construct definition. I then illustrate computer-assisted strategy research that has taken each approach to construct definition, thereby demonstrating that the method of measurement does not preclude either one.

I will distinguish between two approaches defining a strategy as a theoretical construct (approaches explained in Chapelle, forthcoming). The first is a *trait-oriented* definition that conceptualizes a strategy as an attribute of an individual independent of the context in which it is observed. One thinks of a strategy as a trait when one talks about "monitoring" as something learners do all of the time regardless of whether they are listening to an academic lecture, writing an e-mail message to a colleague, or speaking to a close friend. A trait-oriented construct definition assumes that a researcher is able to generalize the inferences made about a construct on the basis of performance on an assessment (i.e., performance in one context) to inferences about the construct in other contexts. In other words, if an individual is a strong monitor user in a test of monitoring, the trait definition would assume that the individual would also be a strong monitor user in the other contexts, such as instructional settings.

A second and contrasting way of defining a strategy as a theoretical construct is an *interactionalist* definition. This definition presents a strategy as a context-dependent attribute of an individual. From an interactionalist approach, one could not define "monitoring" in a global sense. Instead, one would refer to "monitoring while listening to academic lectures," for example.[3] The definition of the strategy would include the context in which the strategy is used.[4] To interpret results of a test of "monitoring while listening to academic lectures," the researcher would generalize results only to monitoring in this context. In short, an important distinction between the two approaches rests on how far the strategy definition assumes the researchers can generalize the results of strategy assessment.

Both approaches to strategy definition have been used in computer-assisted strategy research. The assessment of "advance preparation" illustrates a trait-oriented definition in computer-assisted strategy research (Jamieson & Chapelle, 1987). The strategy is defined in a general way; even though the definition mentions the word "task," it does not refer to any particular task, implying that the strategy is conceived as one that could apply equally to a linguistic task in any context. Table 2 summarizes the key measurement facets of this strategy.

TABLE 2
Examples of trait and interactionalist approaches to definitions of L2 strategies

	Trait-oriented definition	Interactionalist definition
Example	advance preparation (Jamieson & Chapelle, 1987)	resourcing (Chapelle & Mizuno, 1989)
Definition	"planning for and rehearsing linguistic components necessary to carry out an upcoming linguistic task" (O'Malley, Chamot, Stewner-Manzanares, Kupper, and Russo, 1985)	"use of target language reference materials" (pp. 28-29) in the context of learner-controlled CALL materials (p. 26)
Measurement	the amount of time (to .5 second accuracy) between the time that a prompt was given (in a CALL activity) and the time that the student began to respond (averaged over the number of items that student responded to over the course of the semester).	the frequency of the number of requests for help a student made divided by the number of sentences the student produced in a sentence constructing and editing CALL activity (help=dictionary, semantic/pragmatic facts, and grammar)
Inference	performance was assumed to indicate the degree to which the learner was an "advance preparer"	performance was assumed to indicate the degree to which learners used resourcing within the learner-controlled CALL activity
Use	to investigate the relationship between advance preparation and cognitive style as well as the relationship between advance preparation and subsequent language proficiency.	to investigate the extent to which learners use resourcing in a set of learner-controlled CALL activities for practicing grammar and editing (as a means of evaluating the pedagogical potential of optional help)

Advance preparation was measured through response latency in an instructional setting in which learners were working at their own pace. The inferences made on the basis of summed response latencies were the degree to which the language learners were advance preparers in general, rather than the degree to which they used advance preparation while working on this type of software. Accordingly, the scores for advance preparation were used in this research context to investigate the relationship between this strategy and other variables, which were also defined in a context-independent manner.

The assessment of "resourcing" provides an example of an interactional approach to strategy definition (Chapelle & Mizuno, 1989). It is defined in this research as a learner's use of target language reference materials in learner-controlled CALL materials. The definition is interactionalist because it includes the "learner-controlled CALL materials" as the context to which we wish to generalize. Measurement of the strategy was calculated by tabulating the number of times the learners requested help per unit of activity (defined by construction of one sentence). The inferences made were intended to be limited to contexts of learner-controlled CALL, and the scores were used to evaluate the value of offering learners optional help in learner-controlled CALL.

Justifying Inferences about Strategies from Observed Performance Data

In the research cited above, as in any research, one is ultimately concerned about the validity of the uses that are made of the strategy assessment. Validity of test use rests on justifying the inferences made from observed behaviors. Justifications of inferences about strategies constitute evidence for the construct validity of those inferences. Table 3 summarizes and defines types of construct-validity evidence suggested by Messick (1989).

To examine these types of construct-validity evidence, I return to the examples of advance preparation and resourcing. Then I will suggest for each example additional evidence that could be used to make a stronger case for construct validity. Finally, I will underscore two fundamental points about construct validity: (1) the nature of the evidence depends on the way "strategy" is defined in the research and (2) construct-validity evidence refers to the justifications provided for interpretations and therefore it can be evaluated as a strong or weak relative to particular inferences rather than as an all or nothing quality of an assessment.

In the first study, Jamieson and Chapelle (1987) provided three types of justifications that might be used to argue for inferring "advance preparation" from response-latency data. First, content-validity evidence consisted of the authors' judgment that this behavior in the instructional context logically fit with the definition of the construct:

> [A]dvance preparation was inferred from the amount of time it took for the student to press the first key of his or her answer. O'Malley et al. (1985) defined advance preparation as a metacognitive strategy that means "planning for and rehearsing linguistic components necessary for an upcoming language task" (p. 33). The student behavior of consistently waiting before answering may indicate the degree to which he or she was engaging in preparation to answer (p. 531).

TABLE 3
*Potential methods for justifying construct-validity of inferences from tests/assessments (based on Messick, 1989)**

Content analysis	Experts' judgments of what they believe a test measures
Empirical item investigation	Investigation of factors affecting item difficulty and discrimination to provide statistical evidence relevant to researchers' understanding of what the test measures.
Internal consistency	Estimation of the consistency of learners' performance
Empirical task analysis	Documentation of the metacognitive strategies that learners use as they complete test tasks through qualitative methods such as think-alouds.
Correlations with other tests and behaviors	Identification of theoretically predicted levels of covariance among tests and behavior in appropriate contexts.
Experimental studies	Investigation of changes in performance which accompany systematic changes in test design.

* See Chapelle (1994) for an example of how these approaches to construct validity are applied to the evaluation of inferences and uses made from a test.

Second, the authors found a sufficient degree of consistency in strategy use over a six-week period to provide evidence that a construct was assessed through the observed data. "Mean working-style scores from six randomly selected weeks were treated like 6-item scales on which ... reliability estimates were made" (α=.72 and .82 on the two activities, p. 535). Third, a theoretically predicted correlation with a style variable "reflectivity-impulsivity" was found. Advance preparation was significantly, positively related to reflectivity (r=.50; p<.001); one would expect that this strategy "would logically be associated with the slow, careful aspect of the reflective learner" (p. 538).

These three arguments provide some justification for the advance preparation inferences made from learner performance data. However, the argument would be stronger if additional sources of evidence had been provided. First, ideally, evidence consisting of learner verbal reports might indicate that they were thinking about and planning what they would type between the time they heard the input and the time they began to respond. Second, the authors might also have provided correlations not only with a measure of a related construct but also with another measure of the same construct, "advance preparation." Because of the trait-oriented definition of advance preparation, the other measure should assess advance preparation through a different method of measurement to ensure that performance is the same regardless of the context in

which it is measured. Third, some form of experimental data could also contribute to the validity argument. For example, an experiment might compare response latency data of subjects who had been trained to stop and plan with those who were told not to think before responding. If performance reflected the expected patterns for the two groups, results could be used as justification for making inferences about advance preparation from performance.

In the second example, the justification for Chapelle and Mizuno's (1989) use of performance data to assess resourcing rested solely on content evidence for validity. The authors used the following justification: "The computer provides help only upon request so learners must ask for the help they need when they need it. Learners' requests for help are [therefore] evidence of their use of resourcing..." (p. 28). This provides only one argument for their strategy interpretations.

Other justifications that would have made their argument stronger would include the following: First, they might have consulted learners' verbal reports indicating that they had chosen help in order to obtain assistance with the sentence-construction task, rather than for other purposes (e.g., to see what the help looked like). Second, they could have demonstrated consistency in the use of help over the several weeks the activity was used. Third, they might have supplied correlations between use of help on their learner-controlled CALL activity and another one. To act as correlational validity evidence for the interactionalist definition of the strategy, the covariate must be similar to the original assessment in terms of assessing resourcing in learner-controlled CALL as well. Fourth, they could have used an experimental study comparing subjects who had been trained to use help with those who were not told to use it.

These examples indicate that the nature of the construct-validity evidence is related to the construct definition, and where correlational evidence is concerned the type of construct definition (trait-oriented vs. interactionalist) impacts the choice of covariate in validation research. Further, they also indicate that the legitimacy of inferences made from the data is a matter of degree rather than an all or nothing proposition. Validity justification consists of an argument relevant to understanding the meaning of observed data for making inferences about strategies—an argument that is essential for justifying the use of these assessments in SLA research.

Applications of Computer-Assisted Strategy Research

Although very few research projects have used CASA, the method has shown promise in several ways. First, CASA has been useful in extending the researcher's ability to document behaviors in language-learning contexts particularly when large amounts of precise data must be tabulated to make valid inferences about strategies. For years, classroom researchers attempting to study learner strategies have been frustrated by the amount of relevant performance data that they were able to obtain through observation of behaviors (e.g., Naiman, Frohlich, Stern, & Todesco, 1978). Consequently, computer-assisted assessment offers a useful addition to strategy assessment methods.

Second, CASA allows for gathering strategy data during actual instructional exchanges. Given our increasing appreciation of the effects of learners' perception of the task on their performance (e.g., Bruner, 1990), researchers increasingly value data that are obtained in genuine instructional contexts. Despite the usefulness of data obtained through obtrusive methods such as think-aloud protocols (Ericsson & Simon, 1984), and retrospective self-reports (Cohen & Hosenfeld, 1981), there is a need to complement them using observation programs that document behavior unobtrusively in classroom settings. Thus, researchers begin to integrate SLA and pedagogical research and strengthen the possibility of identifying strategies that might provide useful feedback to learners (Scott & New, 1994).

Problems with Computer-Assisted Strategy Research

I would characterize the current problems of computer-assisted strategy research as consisting of two types: analytic measurement problems and practical problems. Measurement problems refer to the need to investigate validity from the perspectives of both construct inferences and consequences of CASA use.

Methods for justifying construct validity are outlined above. As I illustrated, the validity of the inferences and uses of assessment needs to be justified. Because of the precision, accuracy, and directness of data collection in computer-assisted strategy research, however, researchers might attempt to argue away the need for construct validity justification as Chapelle and Mizuno (1989) did: "The behavior exhibited for [resourcing] provides unequivocal evidence for students' use of [this strategy]; in a sense, the behavior is the strategy" (p. 34). From the perspective of the interactionalist construct definition for the resourcing strategy, the behavior is not equivalent to the strategy itself. A second construct validity problem arises when justifications are inappropriate for the type of inferences and uses made from the assessments. For example, if Chapelle and Mizuno (1989) had, on the basis of their content-validity argument, used their data to make inferences about the extent to which learners were resourcers across contexts and had used that data to decide who needed training in resourcing, these inferences and uses would have been inappropriate.

Consequential validity refers to justifications for the usefulness of an assessment for its intended purposes as well as for its unintended outcomes beyond the immediate assessment event and context. Potential consequential validity problems arise when learners are disturbed, rather than facilitated through the assessment process. For example, data gathered and used in a way that violates learners' rights to privacy would argue against consequential validity, as would computer-assisted assessment of learners who feel uncomfortable using the computer. The potential detrimental effects of computer-assisted language learning in general have not been explored rigorously; however, studies in critical pedagogy (e.g., Bowers, 1988) provide some useful directions that may also pertain to the consequential validity of strategy assessment.[5]

Practical problems in CASA are also worth noting. First, in constructing a computer-assisted assessment, researchers may find it difficult to identify software

that simultaneously provides relevant language learning activities for instruction and strategy assessment. Unfortunately, to date little work has attempted to combine efforts in instructional design with those of assessment of either strategies or language. A second practical hurdle for CASA is the challenge of modifying software to get it to gather the appropriate data. Some of the commercial software contains data collection capabilities, but there is no guarantee that a given piece of software will collect the data of interest to the researcher. Many of the CASA projects to date have been conducted by researchers who developed their own software. A third practical problem becomes the management of the large quantities of process data that can be generated by recording the details of learners' interaction. Because disk space is limited, there is always a need to make rational decisions about how to summarize and store the data throughout the assessment process (Goodfellow & Laurillard, 1994). Past research has shown that while these practical problems present significant challenges, they are not insurmountable.

Conclusion

Research on interaction-observation programs has been found useful for assessing some SLA strategies. Methodologically, it is particularly interesting that the researcher can construct the type of learning environments which learners would use routinely for instructional activities but which simultaneously serve as a laboratory for data collection (Doughty, 1992). Moreover, the capability to investigate longitudinally learners' routine "working styles" (Jamieson & Chapelle, 1987) offers an ideal setting for investigating important questions about learners' strategies. For example, how do learners' strategies change as task demands are manipulated or as they accrue experience with a task? How consistent (reliable) is strategy measurement on the same task across different occasions? Can the accuracy of self-report data be substantiated by observation of computer-documented protocols? The investigation of these and other strategy questions, however, relies on the validity of the measurement used to assess strategies and therefore rests in part on validity issues.

Notes

[1] For years, psychologists (e.g., Sternberg, 1977; Snow, 1981) have used response-latency measures for assessing psychological processes in laboratory experiments. In second language acquisition similar research questions have been raised particularly by those approaching SLA from an information processing perspective (McLaughlin, 1987). Automaticity in language processing is hypothesized to be indicative of language knowledge efficiently stored for expedient retrieval; as an aspect of the target language is better learned, restructuring of knowledge occurs making access more automatic. In experimental settings, the amount of time subjects take to respond to a task has been used as an indicator of how automatic subjects are in the use of the linguistic knowledge necessary for performance on the task, or whether knowledge is implicitly

or explicitly stored. For example, Hagen (1994) used a computer program to present items to subjects and to time their responses to make such inferences. Crucial to the interpretation of time-to-respond in this case was the learners' understanding that they were to respond as quickly as they could—a task demand the researcher had to make clear to the subjects.

[2] This is not the only way of viewing a strategy. Another is to define it in terms of the actual observed behaviors (see Cohen, this volume).

[3] Of course, the context of the academic lecture could be defined in greater detail to state a more specific construct definition.

[4] What is needed to better express the interactionalist construct definition is a more complete and systematic way of defining "context." Proposals for approaching this problem have been suggested by researchers in language testing (Bachman, 1990), language instruction (Skehan, 1992), and SLA research (Duff, 1993).

[5] Bowers (1988) points out the need to examine the range of possible activities learners engage in through computer-assisted instruction in order to shed light on what students are missing by spending time at the computer. She also guides us to examine the cultural ideologies—such as the value Western societies place on information and individualization—inherent in our educational uses of technology.

References

Abraham, R. A. (1996). Using task products to assess second language learning processes. *Applied Language Learning, 7*(1).

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *Modern Language Journal, 65*, 24-35.

Bland, S. K., Noblitt, J. S., Armington, S., & Gay, G. (1990). The naive lexical hypothesis: Evidence from computer-assisted language learning. *Modern Language Journal, 74*, 440-450.

Bowers, C. A. (1988). *The cultural dimensions of educational computing.* New York: Teachers College Press.

Bruner, C. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University Press.

Chapelle, C. A. (forthcoming). Construct definition and validity inquiry in SLA research. In L. Bachman & A. Cohen (Eds.), *Second language acquisition and language testing interfaces.*

Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research, 10*(2), 157-187.

Chapelle, C., & Mizuno, S. (1989). Students' strategies with learner-controlled CALL. *CALICO Journal, 7*(2), 25-47.

Chun, D. M. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System, 22*(1), 17-31.

Cohen, A. D. (1996). Verbal reports as a source of insights into second-language learner strategies. *Applied Language Learning, 7*(1).

Cohen, A.D., & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. *Language Learning, 31*, 285-313.

Doughty, C. (1992). Computer applications in second language acquisition research: Design, description, and discovery. In M. Pennington, and V. Stevens (Eds.), *Computers in applied linguistics: An international perspective* (pp. 127-154). Clevedon, Avon: Multilingual Matters, Ltd.

Doughty, C. (1987). Relating second-language acquisition theory to CALL research and application. In W.F. Smith (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 133-167). Lincolnwood, IL: National Textbook.

Doughty, C., & Fought, C. (1984). On investigating variable learner response: Toward achieving better CALL courseware design. Report from the language analysis project. Philadelphia: University of Pennsylvania. (cited in Doughty, 1987)

Duff, P. (1993). Tasks and interlanguage performance: An SLA perspective. In G. Crookes & S. M. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp. 57-95). Philadelphia: Multilingual Matters.

Ericsson, K., & Simon, H. (1984). *Protocol analysis—Verbal reports as data*. Cambridge, MA: The MIT Press.

Goodfellow, R., & Laurillard, D. (1994). Modeling learning processes in lexical CALL. *CALICO Journal, 11*(3), 19-46.

Hagen, L. K. (1994). Constructs and measurement in parameter models of second language acquisition. In E. E. Tarone, S. M. Gass, & A. D. Cohen (Eds.). *Research methodology in second-language acquisition* (pp. 61-87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Higgins, J., & Johns, T. (1984). *Computers and language learning*. New York: Addison-Wesley.

Hsu, J., Chapelle, C. &, Thompson, A. (1993). Exploratory environments: What are they and do students explore? *Journal of Educational Computing Research, 9*(1), 1-15.

Hulstijn, J. (1993). When do foreign language learners look up the meaning of unfamiliar words? The influence of task and learner variables. *Modern Language Journal, 77*(2), 139-147.

Jamieson, J., & Chapelle, C. (1987). Working styles on computers as evidence of second language learning strategies. *Language Learning, 37*, 523-544.

Krashen, S. (1982). *Principles and practices in second language acquisition*. Elmsford, NY: Pergamon Press.

McLaughlin, B. (1987). *Theories of second language acquisition*. London: Edward Arnold.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement (3rd ed.)*. (pp. 13-103). New York: Macmillan.

Naiman, N., Frohlich, M., Stern, H., & Todesco, A. (1978). *The good language learner. Research in education series (Report No. 7)*. Toronto: Ontario Institute for Studies in Education.

O'Malley, J., Chamot, A., Stewner-Manzanares, G., Kupper, L., & Russo, R. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*, 21-46.

Skehan, P. (1992). Second language acquisition strategies and task-based learning. *Thames Valley University Working Papers in English Language Teaching, 1*, 178-208.

Scott, V. M. & New, E. (1994). Computer aided analysis of foreign language writing process. *CALICO Journal, 11*(3), 5-18.

Snow, R. E. (1981). Toward a theory of aptitude for learning: Fluid and crystallized abilities and their correlates. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 345-362). New York: Plenum.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, New Jersey: Lawrence Erlbaum.

Wenden, A. (1985). Learner strategies. *TESOL Newsletter, 19*, 1-7.

Acknowledgment

I am grateful for comments from Roberta Abraham and an anonymous reviewer on earlier versions of this article.

Author

CAROL A. CHAPELLE, Professor, TESL/Applied Linguistics, 203 Ross Hall, Iowa State University, Ames, IA 50011. Specializations: computer-assisted language learning, language testing, SLA research.

Applied Language Learning
1996, Vol. 7, Nos.1 & 2, pp. 61-89

Using Task Products to Assess Second Language Learning Processes

Roberta G. Abraham
Roberta J. Vann
Iowa State University

Since the shift of attention in the 1970s to the processes language learners employ in acquiring second language (L2) proficiency, the products of performance (language or other representation of meaning that a learner formulates in response to a task) have often been ignored. However, as Bialystok (1990a) demonstrates with examples from research on communication strategies, when products are examined within appropriate conceptual frameworks that reflect actual processing options, considerable information about underlying intentions and processes is revealed. This article reviews Bialystok's analysis of the communication strategy work and then describes two studies in which products are interpreted within frameworks that illuminate underlying processes of L2 learning. In the first study, reported by Abraham and Vann (1994), the responses from a cloze test are examined within Bialystok and Ryan's (1985) framework of analyzed knowledge and control to reveal L2 learners' metacognitive skills. In the second, carried out by Vann and Schmidt (Schmidt & Vann, 1992; Vann & Schmidt, 1993), a depth of processing framework derived from work by Craik and Lockhart (1972), Marton and Saljo (1984), and others guides the analysis of L2 learners' reading notes to reveal the strategies the learners use as they interact with an academic text. Important to the discussion are arguments for the validity of the interpretations and proposed uses of the cloze responses and reading notes. We conclude that products, appropriately analyzed, can provide important insights into the processes of L2 learners.

In the late 1970s, an important shift occurred in L2 research, away from interest in the *products* of language learning towards the *processes* learners employ in acquiring the knowledge and skills required of a proficient L2 user. However, even then, a cautionary note was sounded by a group of researchers who were committed to discovering the processes, also referred to as "strategies and techniques" (Naiman, Frohlich, Stern, & Todesco, 1978, p. 65), that learners employed. In their exemplary report on a well-designed and complex study, they noted that while "the practice of inferring language learning *processes* solely on the basis of language *product* is tenuous" (p. 65,

© 1996, Roberta G. Abraham and Roberta J. Vann

959

italics in original), in future research a combination of process and product data would be needed to understand how second language learners develop proficiency. By "products," they meant learners' performance on individual exercises and learning tasks, which we take in this paper to include the language or other representation of meaning (e.g., diagrams, sketches, or gestures) that a learner formulates in response to a task (e.g., completing a cloze exercise, writing a composition, taking notes on a reading passage or a lecture, or describing a series of pictures that tell a story).[1]

However, since 1978, relatively few second language acquisition (SLA) researchers have made use of products to identify and understand the processes and strategies[2] that learners use in performing a particular task. Many researchers appear to perceive process and product as dichotomous (product *versus* process), viewing products as irrelevant in investigating learner strategies. Our purpose here is to dispel this notion by reviewing literature discussed in detail by Bialystok (1990a) in which products have been used to discover the communication strategies of L2 learners, and then by describing two of our own studies in which products were important in illuminating L2 learning processes. An important part of this discussion addresses the validity of the inferences drawn from the product data.

Products as Indicators of Process in L2 Communication

The products most frequently used to shed light on L2 communication strategies have been learners' descriptions of drawings of objects and complex pictures (Tarone, 1977; Varadi, 1980; Bialystok & Frohlich, 1980; Bialystok, 1990a). Other products have resulted from picture reconstruction (Bialystok, 1983), translation (Galvan & Campbell, 1979; Varadi, 1980), sentence completion (Blum-Kulka & Levenston, 1983), conversation (Haastrup & Phillipson, 1983), narration (Dechert, 1983; Raupach, 1983), instruction (Wagner, 1983), word transmission (Paribakht, 1985), and interview (Raupach, 1983). In an extended large-scale study referred to as the Nijmegen project (Bongaerts, Kellerman, & Bentlage, 1987; Kellerman, Bongaerts, & Poulisse, 1987; Poulisse, 1987; Kellerman, Ammerlaan, Bongaerts, & Poulisse, 1990), the products from several types of tasks were examined. These tasks included describing concrete objects and abstract shapes, retelling stories, and interviewing.

In some cases, utterances from these products were categorized in taxonomies such as that developed by Tarone (1977).[3] While these taxonomies have been useful in identifying actual behaviors, Bialystok (1990a) reports a study that demonstrates that they are not without problems of reliability and validity. Reliability concerns, of course, arise in the coding of strategies. Using Tarone's taxonomy, Bialystok found that a single utterance might reflect several strategies, raising the question of whether to code each separately or to evaluate the utterance globally. Further, although Tarone's definitions of categories were explicit and her examples clear, a number of ambiguities became apparent in the coding of individual utterances. As a result of these difficulties, Bialystok estimated that in a reliability check, disagreements between coders occurred for at least 50% of the utterances.

Validity problems were encountered in associating a linguistic utterance provided by the learner with the underlying intentions and processes that produced it. As Bialystok (1990a) puts it, "[i]f the taxonomic descriptions are valid, then the distinctions should correspond to real alternatives or real choices experienced (at some level although not necessarily consciously) by the learner" (p. 56), that is, the choice of utterance should be related systematically to some characteristic of the learner or the environment. However, research attempting to link particular types of utterances to learner or situational factors, for example, learner proficiency (Tarone, 1977; Bialystok, 1983; Paribakht, 1985), elicitation task (Bialystok & Frohlich, 1980), and first language (Tarone, 1977), that might affect these choices has produced few meaningful results. Along these same lines, Bialystok (1990a) asked L2 learners to use descriptions selected as clear representatives of categories in Tarone's taxonomy in identifying objects from an array. She found little relationship between category of strategy and accurate identification of an object by the listener.[4]

Overall, the taxonomic approach to the study of communication strategies, while yielding valuable insights into the types of surface expressions learners use in coping with communication difficulties, has failed to capture the essence of learner choice and intent, and in this sense, lacks validity. To illustrate how one might probe more deeply into such questions, Bialystok describes two large research programs in which the investigators, rather than simply allowing strategic categories to emerge from the data, selected an appropriate conceptual framework within which to interpret the products so that they revealed the options available to the participants. In the first, Snow, Cancini, Gonzalez, and Shriberg (1989a, 1989b) examined definitions of common objects by bilingual children at the United Nations International School in New York within a framework relating children's use of "decontextualized" metalinguistic use of language to "reading skills and 'literacy' in general" (Snow et al., 1989b, p. 5, quoted in Bialystok, 1990a, p. 105). The definitions were classified as "formal" (containing an equivalence and a superordinate, e.g., "a donkey is an animal") and "informal" (failing to specify such semantic relationships). The ability to produce formal definitions was considered to be an important indicator of metalinguistic development because of the "explicit and 'intentional' use of words outside of the contexts which endow them with meaning" (Bialystok, 1990a, p. 105). While Snow and her colleagues quantified the additional descriptors that contributed to the quality and communicative adequacy of the definitions in order to relate them systematically to the children's progress in achieving literacy, the aspect of their work of relevance here was the identification of factors that did (and did not) predict that a definition would be formal. Academic language proficiency, as measured by reading and language scores on the California Achievement Test (Snow et al., 1989a), predicted the use of a formal definition; for children tested in both L1 and L2, the language used did not, nor did oral proficiency in the L2, though the latter was related to the quality of the definition (Bialystok, 1990a). Because the formal and informal categories of response, "rooted in a conception of what these options mean for language processing" (Bialystok, 1990a, p. 109), make a distinction that could be empirically related to general development of

a learner, they are valid in the sense described above.

The second study described by Bialystok was the Nijmegen project, in which the referential (lexical) communication strategies used by Dutch learners of English in L1 and L2 on several tasks were observed. Again, to avoid the weaknesses of the taxonomy approach, the "tendency to confuse the linguistic realization of the referential strategy with the strategy itself"[5] and the confusion of the strategy with the properties of the referent [6] (Kellerman et al., 1990, p. 165), the researchers used a framework based on a theoretical view of communication and language production. In this view, learners unable to communicate their ideas can either "manipulate the concept so that it becomes expressible through their available linguistic . . . resources, or . . . manipulate the language so as to come as close as possible to expressing their original intention" (Kellerman, source unknown, quoted in Bialystok, 1990a, p. 111). Accordingly, the researchers used only two categories in coding the data: *conceptual*, in which the speaker "analyses the concept by decomposing it into its criterial features" either by listing them or referring to a related concept, and *linguistic*, in which the speaker "manipulates his linguistic knowledge" by strategies of "morphological creativity" or transfer from the first language (Poulisse, 1987, pp. 146-147). Again, while the researchers used products as data, their choice of categories was process-oriented, and the results were subsequently related to learner and situational factors. As in the Snow study, the analysis showed that, in most cases, neither the language used in the response (Dutch or English) nor the learners' proficiency predicted the type of strategy chosen (though the quality and effectiveness of the strategy did vary with proficiency). However, the nature of the task did seem to influence the choice of strategy, with descriptions of single objects eliciting a relatively low proportion of "code" strategies (the "linguistic" category expanded to include certain "ostensive" strategies), story-retellings evoking proportionally more, and interviews producing the highest number (Bialystok, 1990a, p. 113), thus providing evidence for the validity of the classification scheme. Bialystok (1990a) further validated the two-category system by reanalyzing the taxonomy data in her picture identification task within the framework used in the Nijmegen project. She found a distinct difference in the effectiveness of communication for these two strategy types, with those characterized as "code" inducing a higher percentage of correct listener identifications than those characterized as "conceptual."

The work reported here on the use of products to shed light on the communication strategies of L2 learners has important implications for those interested in L2 learning processes. Products can be useful in inferring these processes, but only when the products are analyzed within a conceptual framework that permits their reliable classification and valid interpretation. That is, the framework must allow for and distinguish among different stages of learner development and/or options available to the learner, and the researcher must be able logically to relate the products to these stages or options. The communication strategy work suggests that there will likely be only a few broad product categories growing out of a conceptual framework, but that each of these may include several manifestations of the underlying process (e.g., in the Nijmegen project, the broad "code" and "linguistic" categories each included more than one type

of utterance). However, the framework allows the researcher to interpret these manifestations as the result of a learner's stage of development or his/her exercising a single underlying option (e.g., deciding to amend her language to make herself understood), and thus avoids the validity problems encountered in interpreting taxonomy results where the categories simply emerged from the data.

In the remainder of this article, we describe two studies of L2 learning processes in which we make use of the insights from the communication strategy work in interpreting task products. In the first, we examine cloze responses to discover the nature of learners' metacognitive skills; in the second, we analyze reading notes to assess the strategies employed as L2 learners interact with academic texts. For each study, we discuss the selection of a framework within which to analyze and interpret the data, the distinctions in the framework that underlie the classification of products, the analysis and coding procedures, and the findings. On the basis of this discussion, we argue for the validity of our interpretations and proposed uses of these products, and suggest that products, appropriately analyzed, can illuminate the processes of L2 learners.

Study One
Cloze Responses as Evidence of Metacognitive Skills of ESL Learners

This study, reported by Abraham and Vann (1994), examined the metacognitive skills displayed on a cloze task administered to a group of Lebanese ESL learners several months after they entered the beginning level of an intensive English program. Although these learners had no prior knowledge of English when they entered the program, their general English proficiency varied considerably at the time the cloze was administered. The major purpose of the study was to compare the metacognitive skills of learners who were successful on the cloze with those of learners who were not. While think-aloud procedures were used to collect data, one goal of the study was to discover how much valid information about metacognition could be gained from analysis of the cloze responses, or *products* of the task. This report focuses on insights gained from the product data.

Background

Although "metacognition" has been defined in many ways and investigated from many points of view in the past 25 years (for example, see Flavell & Wellman, 1977; Brown, Bransford, Ferrara, & Campione, 1983; Yussen, 1985), researchers and educators agree that this "knowing about knowing" (Garner & Alexander, 1989, p. 143) or "thinking about thinking" (Yussen, 1985, p. 253) or "executive control" (Brown et al., 1983, p. 110) plays an important role in learning. Although much of the research on metacognition has dealt with learning and reading in the learner's native language (for example, see reviews in Brown et al., 1983, and Garner, 1987), metacognition has not gone unnoticed in L2 research. Wenden (1987) and Reid (1987) document L2

learners' metacognitive knowledge of their own abilities and preferences for learning. Horwitz (1987) describes research using a questionnaire to assess learner beliefs about learning. Jamieson and Chapelle (1987) investigated learners' use of several executive control strategies by means of computer-assisted techniques. O'Malley and Chamot (1990) identified from interview and think-aloud data seven aspects of executive control (planning, directed attention, selective attention, self-management, self-monitoring, problem identification, and self-evaluation) used by foreign language learners. Oxford (1990) advocates the teaching of three types of metacognitive strategies: centering, arranging and planning, and evaluating one's learning. Vann and Abraham (1990) noted the similarity in the cognitive strategies used by successful and unsuccessful learners and pointed to the need for metacognitive explanations for the observed differences in proficiency. Bachman (1991) describes a framework for L2 test development in which one aspect of language ability to be measured is strategic competence, composed of the metacognitive strategies of assessment, goal-setting, and planning, and Block (1992) compared the metacognitive processes of comprehension monitoring in first- and second-language readers.

Most of the L2 research on metacognition has focused on self-understanding and control, that is, what learners know about how they learn and how they oversee and regulate their mental processes as they learn or perform tasks requiring application of what they have learned. Although this work has provided valuable insights into learner differences, its focus neglects an aspect of cognition that seems highly relevant in our quest to understand why some learners are successful and others are not. What is missing in the "control" approach to metacognition is a consideration of the knowledge of language that learners develop and manipulate in learning or performing a task. Metacognitive strategies are not used in a vacuum; learners may be adept at planning, monitoring, and evaluating, but without a firm knowledge base to work from, these metacognitive strategies will not produce success.

In analyzing the data from this study, we used a theoretical framework of metacognition that takes into account both knowledge and control. The framework, proposed by Bialystok and Ryan (1985) to provide insight into broad issues in both first- and second-language development, differs from those used in most discussions of metacognition in that it comprises two "theoretically orthogonal" (p. 208) but pragmatically related dimensions, analyzed knowledge and control,[7] which can be used to characterize both task requirements and learner abilities. Along the first dimension, knowledge of the language ranges from *unanalyzed* information, which is used routinely, with little or no learner awareness of its structure, to *analyzed* information, which the learner is able to manipulate and use creatively. Bialystok and Ryan illustrate this dimension with the example of subject-verb agreement. Some learners simply produce sentences that contain such agreement without understanding the principles involved, thus demonstrating unanalyzed knowledge. Others have access to structural concepts of predicate and arguments, which permit "construction and manipulation" (p. 211) of sentence parts (e.g., correcting sentences with faulty agreement or stating the rules for agreement), thus demonstrating analyzed knowledge. Cloze

tasks such as the one used in this study require a relatively large amount of analyzed knowledge, and learner variation along this dimension can be documented.

The control dimension, as in much of the earlier work on metacognition, "represents an executive function that is responsible for selecting and coordinating the required information within given time constraints" (Bialystok & Ryan, 1985, p. 213). Again, tasks such as the cloze, which require learners to attend to and coordinate multiple characteristics of structure and meaning in order to select appropriate lexical items, make high demands on learners' control abilities and are appropriate in demonstrating individual differences in this respect.

Applying the Bialystok and Ryan framework to the Lebanese study discussed here, we examined not only learners' use of control strategies (including those of predicting, planning, monitoring, and evaluating described by other researchers) but also the nature of their knowledge of English accessed in the process of completing the cloze task (was that knowledge sufficiently analyzed for them to fill each blank with a word that met all the syntactic and semantic constraints of its context?). We expected that learners who performed well on the cloze would show a high degree of control and analyzed knowledge, whereas those who performed less well would show deficiencies in at least one dimension of metacognition, deficiencies that would likely be reflected in their scores on other measures of academic language proficiency such as the TOEFL. The degree to which these expectations were met provides a measure of the validity of using cloze products to reveal L2 learners' metacognitive processes.

Method

Participants

The participants were nine male Lebanese students who, under the sponsorship of the Hariri Foundation, entered the beginning level of the academically oriented Intensive English and Orientation Program (IEOP) at Iowa State University. These scholars had been carefully selected as intelligent, hardworking, and highly motivated to succeed in their respective areas of academic specialization. All were native speakers of Arabic, and none had studied English prior to their arrival in the U.S. Although their scores on the Michigan English Proficiency Test placed all of them into the lowest level (of five) in IEOP, differences in their abilities began to emerge as early as midterm, and by the end of 24 weeks of instruction, their TOEFL scores ranged from 377 to 510.

Procedure

The cloze task was the third of four tasks in which the Lebanese learners were asked to think aloud as they worked. It was administered four to six months after the learners entered the program. Taken from Cohen's (1980) *Testing Language Ability in the Classroom*, the 33-blank cloze passage dealt with cross-cultural differences, presumably a topic with which the Lebanese learners had firsthand experience. As in most clozes, the deleted words differed along a number of dimensions (e.g., whether

they were content or function, how much context was needed to restore them, whether they occurred elsewhere in the passage) that affect difficulty (Brown, 1988; Abraham & Chapelle, 1992). A number of items proved to be quite difficult for the learners in this study, thus eliciting the metacognitive skills of interest.

As the learners worked, they were asked to think aloud in the presence of a research assistant, a female Chinese graduate student in TESL.[8] The research assistant's role was limited to probing learners' thoughts when they were not being expressed and answering questions on procedures or vocabulary.

Data Analysis

The analysis of product data consisted of examining the responses to the cloze to identify patterns that could be associated with the metacognitive skills of learners who were successful on the cloze and those who were not. The analysis was organized around three questions, the first focusing primarily on one aspect of control (predicting and planning, in this case selecting appropriate cognitive strategies for the task), and the last two on aspects of both analyzed knowledge and control (selecting and coordinating the required information). The questions and the ways in which products were involved in providing answers are shown below.

1. Did learners appropriately identify the cognitive strategies needed for the task? One important cognitive strategy needed for the task was identifying the main ideas of the passage. Because the cloze proved difficult for these learners, it was assumed that those who demonstrated understanding of the main ideas must have recognized the need to do so. To assess understanding, the researchers individually examined each learner's responses to the key blanks (those carrying important content information) in each of the six main ideas in the passage and rated each idea as "not understood," "partially understood," or "understood." Disagreements were jointly resolved by the researchers.

Another cognitive strategy that could be inferred from products was avoiding guesses that would be misleading in filling in later blanks. Since learners were given no instructions that discouraged them from guessing, unfilled blanks were assumed to reflect the use of this strategy.

2. Did learners demonstrate adequate knowledge of the lexicon and access it appropriately in filling in blanks? The number of correct responses gave one answer to this question, but incorrect responses provided clues concerning learners' analyzed knowledge of the semantic features of lexical items and their ability to access appropriate words for particular contexts. Thus, incorrect answers were classified according to the following criteria: Were they semantically appropriate for the context? Were they in the general semantic field of the passage, though not appropriate in the specific context? Were they totally inappropriate semantically (or too vague to make a judgment)? Were they the opposite of the correct answer? Discrepancies in coding were again jointly resolved.

3. Did learners demonstrate adequate morphological and grammatical knowledge of English and access it appropriately in filling in blanks? As for question 2, both

the number of correct responses and the nature of incorrect responses were useful in assessing learners' analyzed knowledge and control. Here, the analyzed knowledge of interest was morphosyntactic, and the researchers individually coded each incorrect answer as syntactically appropriate for the entire sentence, syntactically appropriate if only part of the sentence was considered (operationally, whether the response could fit syntactically with at least one word adjacent to the response), or totally inappropriate syntactically in the sentence or any string of words within it. Again, the researchers jointly resolved discrepancies in coding.

Results

Scores on the cloze ranged from a high of 20 (out of a possible 33) to a low of 6. As expected, the patterning of learners' cloze scores was similar to the patterning of their scores on TOEFL seven months after they entered the program, i.e., Ahmed,[9] Bashir, Khalid, and Daaboul had the highest scores on both tests, whereas Fouad, Ghassan, Hossein, and Jahad had the lowest scores on both (see Table 1).

TABLE 1
Cloze and TOEFL Scores

Subject	Cloze Score	TOEFL Scores after 24 weeks of instruction[a]
Ahmed	20	483 (487)
Bashir	19	477 (510)
Khalid	16	443 (437)
Daaboul	15	510
Essam	13	423
Fouad	10	407
Ghassan	10	377
Hossein	8	400 (420)
Jahad	6	383

[a] First scores shown are for TOEFL administered immediately after instruction was complete. Scores in parentheses are for TOEFL administered three weeks later.

To sharpen the distinction between successful and unsuccessful performers on the cloze, the following analyses contrast the responses of the two highest scorers (with scores of 20 and 19) with those of the four lowest (with scores from 10 to 6). The results of the various product analyses described above are summarized in Table 2. Of interest here are the qualitative differences between the successful and unsuccessful learners on four key variables that provide evidence of differences in metacognitive

TABLE 2
Performance of successful and unsuccessful learners

					Incorrect Responses						
					Semantic				Syntactic		
Learner	Score	Ideas	Omits	Incorrect	Approp.	Tot. inapprop.	Gen. sem. field	Opp.	Gram.	Tot. ungram.	Too small cntxt.
Successful											
Ahmed	20	4.5	5	8	2	0	6	0	6	0	2
Bashir	19	4.5	12	2	1	0	1	0	2	0	0
Unsuccessful											
Fouad	10	2	0	23	3	9	8	3	8	1	14
Ghassan	10	3.5	0	23	7	7	7	2	7	1	15
Hossein	8	1.5	1	24	3	15	5	1	7	3	14
Jahad	6	1	3	24	3	13	8	0	5	2	17

Ideas = number of main ideas subject appeared to understand (partially understood idea counted as .5)

Omits = number of blanks left empty

Approp. = number of incorrect answers that were semantically appropriate for the context

Tot. inapprop. = number of incorrect answers that were totally inappropriate semantically or too vague to allow a judgment

Gen. sem. field = number of incorrect answers appropriate to general semantic field of passage but not appropriate in specific context

Opp. = number of incorrect answers that were opposites of the correct answers

Gram. = number of incorrect answers that were syntactically appropriate for the context

Tot. ungram. = number of answers totally inappropriate syntactically in any grammatical unit within the sentence

Too small cntxt. = number of answers that were syntactically inappropriate in entire sentence but formed acceptable sequence with at least one adjacent word

abilities. First, the two most successful learners, Ahmed and Bashir, understood more of the six major ideas (4.5) than any of the unsuccessful learners. Second, the two successful learners, while filling nearly two thirds of the blanks with correct responses,

omitted a substantial number of the remaining blanks, suggesting that they recognized the dangers of filling in words that they were unsure of. In contrast, the unsuccessful learners left very few blanks unfilled.

Third, the semantic analysis shows that the unsuccessful learners used far more words that were totally inappropriate in the context or were related only to the general topic of the passage than the successful learners. This comparison suggests that the unsuccessful learners had some knowledge of the needed vocabulary, but were unable to select words with precisely the features required by the contexts (because they either lacked these words altogether or were unable to access them on this occasion). Finally, the syntactic analysis shows that the unsuccessful learners used far more words that fit in a local context but not in the sentence as a whole. Like the semantic comparison, this finding suggests that the unsuccessful learners' analyzed knowledge (or control) of English morphology and syntax was inadequate to permit the manipulation required by the cloze or that they simply did not understand overall meaning.

In addition to differentiating the successful from the unsuccessful learners, the product analysis points to at least two differences within the unsuccessful group that would have been useful had follow-up intervention been attempted. The first is in understanding the main ideas of the passage, where Ghassan clearly contrasts with the other three, despite his overall low score. In intervention, it would not have seemed necessary to encourage Ghassan to search for meaning, though the others would likely have benefited from training to develop skills for identifying main ideas. The second is in the number of answers that were totally inappropriate semantically, with Fouad and Ghassan producing only slightly over half as many as Hossein and Jahad. These data suggest an unproductive pattern of wild guessing in Hossein's and Jahad's approaches that would likely be observed in other contexts. In intervention these two learners could have been shown ways to make better guesses by drawing on what they did know.

Discussion

Anchored in Bialystok and Ryan's theoretical model of metacognition, the comparisons reported above support the validity of interpreting products as indicators of learners' metacognitive abilities. The patterns of skills so revealed are related to the learners' degree of success on the cloze and point to important strengths and weaknesses of the unsuccessful learners that could be used in intervention. However, these claims could be strengthened, and would perhaps be modified, by more direct evidence of learners' thought processes.[10] For example, for question 1 above, concerning whether learners identified the cognitive strategies needed for the task, an assumption was made that learners who filled in key blanks correctly recognized the need to discover main ideas, and, indeed, the think-alouds of the successful learners confirmed this assumption. However, those who were less successful in completing the cloze

may also have recognized this requirement but have been unable to meet it because of lack of the necessary analyzed knowledge; the think-alouds revealed that this was probably the case for the least successful learner (Jahad). Similarly, the interpretation of unfilled blanks as representing an unwillingness to make potentially misleading guesses could be further validated by learner comments in think-alouds and/or retrospective reports to this effect. Here, the think-alouds of the two successful learners showed that they were intentionally leaving some blanks unfilled because they did not know the answers; additional probing might have revealed why they felt it was desirable to do so.

Revelation of the learners' thought processes would undoubtedly have modified the classification of products in Questions 2 and 3. For example, without knowledge of what the learner was thinking, we classified any word that could *conceivably* fit syntactically with an adjacent word as appropriate in a limited context. The classification of some responses in this category was confirmed in the think-alouds, but in many cases, it was not. Another question that remained unresolved through product analysis was whether an incorrect response resulted from lack of analyzed knowledge or of the control necessary to access the required knowledge. Unfortunately, the think-alouds were of little help in resolving this issue.

Despite these limitations, the analysis of products in this study did illuminate learner processes that would likely remain invisible in other methods of data collection, particularly by showing where learners' guesses were "good" or "bad," and suggesting areas where learners' analyzed knowledge (or control) was deficient. For teachers lacking the time and resources to have learners self-observe or self-report, products from tasks such as the cloze have an obvious advantage in that they are easily available throughout the period of instruction. This study suggests that, appropriately analyzed, these products can provide valuable information pertaining to intervention. The outcomes of such intervention would provide further evidence of the validity of the product analysis. For researchers seeking detailed information about how learners learn, product analysis can suggest processes that can be confirmed through other data collection methods and sometimes provide insights not available elsewhere.

We turn now to another study in which products provided a different type of information about learner processes.

Study Two
Using Reader Notes to Assess Academic Reading Strategies

Though notetaking is viewed as an essential academic learning strategy, learner notes have rarely been used as a means for *assessing* strategies. We suggest here that learner notes may be a product that can help us view important academic reading strategies that a reader uses in interacting with a text.[11] We draw our support for the validity of using learner notes to assess learner strategies primarily from a study conducted by Schmidt and Vann (1992; Vann & Schmidt, 1993) which investigated the characteristics of notes made by L2 learners while reading academic texts and attempted to link

these findings with traits of the learners, including cultural background and level of schooling. In addition to providing evidence that categories of notetaking behavior can be linked with certain influential learner variables, we argue that the naturalness of notetaking enhances its validity in assessment.

Background

Although little direct use has been made of learner notes to assess learner strategies, research has been conducted on the usefulness of notes in enhancing comprehension and learning in first language reading. These studies help reveal the potential of notes as an assessment tool. Notes have been shown to be effective, even when not reviewed, but it is the *quality* of the notes, not merely the act of taking them, that makes a difference in learning (Anderson & Armbruster, 1980). Even when they are not reviewed, notes can enhance retention and comprehension when they focus on text analysis and increase the depth at which material is processed. Anderson and Armbruster (1980) and Bretzing and Kulhavy (1979) suggest that when note-taking seems ineffective it may be because subjects fail to engage in deep and meaningful processing. Further support for the power of certain kinds of notes over others comes from a study done by Hidi and Klaiman (1983) which identified the notetaking strategies of experts and novices, finding that the former tended to "construct certain mental representations" (p. 381), to keep the overall purpose of the text in mind, to perform more critical appraisal of the text, and to paraphrase rather than to copy verbatim. Novices in the study took notes that resembled those made by children (as described by Brown & Smiley, 1978) in relating more closely to the surface structure of the text, both in order of the occurrence of the idea units and in the near verbatim selection of words produced. Depth of processing has also been specifically associated with certain categories of notetaking, especially graphic representations of text, including diagrams or knowledge maps of the content (Jones, Pierce, & Hunter, 1988/89; Tang, 1992/93; and Amer, 1994).

The fact that some kinds of notes may reflect or enable deeper engagement and greater learning connects with a larger body of work relating depth of processing with recall (Craik & Lockhart, 1972; Craik, 1977). This notion that "deep" as opposed to "surface" processes are more likely to involve learning also figures prominently in the work of Marton and his colleagues (Marton & Saljo, 1984), who originally probed (using questionnaires and interviews) what university students learned from reading an academic text and found that different approaches to the task were associated with qualitative differences in outcomes. While some students fully understood the argument being presented by the author, others could mention only details. Those who set out trying to understand the content of the text and engaged themselves deeply in trying to understand it stood the best chance of getting the main ideas and supporting facts. The original study was subsequently extended to a variety of tasks such as writing assignments and scientific problem solving. Learners taking a "deep approach" to a reading task focus on "what is signified" (i.e., the author's argument) and attempt to

taking a surface approach to the task focus on the "signs" (i.e., the text itself and discrete elements such as vocabulary). "Surface" readers depend on memorizing and rote learning, failing to distinguish major from supporting details or new information from old.

Marton and Saljo (1984) suggest that a surface approach to reading results in poor comprehension, while a deep approach makes understanding the author's message more likely. Applying a similar distinction to writing, Bereiter and Scardamalia (1987) distinguish "knowledge-telling," mere regurgitation of ideas, from knowledge which is transformed in some way. Haas (1990) argues that the extent to which knowledge is transformed distinguishes composing from mere transcription, which involves little or no reformulating, and Spivey (1990) maintains that certain kinds of reformulating (organizing, selecting, and connecting) are key operations in constructing meaning from text.

Taken as a whole, this literature, spanning two decades, paints a convincing picture of the links between certain deep processing operations and the construction of textual meaning on the one hand and surface processing and failure to comprehend on the other. However, because it is not clear to what extent learners are aware of their own use of these strategies, assessment via learner report remains problematic. Bretzing and Kulhavy (1979) suggest that this problem can be solved in part if evidence of deep or surface processing can be discerned in the notes learners take, and this line of reasoning underlies the study by Vann and Schmidt described here.

The researchers, both of whom teach ESL for academic purposes, had noticed that their students spontaneously took notes while reading in class, even when not specifically asked to do so. This observation suggested that notetaking was a by-product of academic reading for most ESL readers, making it an unobtrusive means of gathering data during the reading process when compared with methods that require students to read aloud, answer questions periodically while reading, etc. Further, the researchers observed that students varied remarkably in their notetaking, both in what they chose to note and in their style of conveying it. Prior research suggesting the superiority of some notetaking styles over others encouraged the researchers to examine student notes systematically within the conceptual framework of deep and surface processing to determine the extent to which notes might convey insights into how readers were interacting with texts.

The researchers thus set out to assess certain academic reading strategies of advanced students in an intensive English program. They were interested in whether notetaking strategies, classified according to the deep/surface processing dichotomy, were predictably related to the context of reading, for example, to different texts or tasks, and whether or not cohorts of strategy users were apparent. For example, did students with similar overall English language proficiency, similar amounts of academic training, and/or related first languages show the same tendencies in notetaking? A basic question in the study, on which we focus here, is the extent to which the results provide evidence for the claim that notes are a valid means of assessing academic reading strategies.

Method

Participants

The participants were 50 students (27 women and 23 men) enrolled in the advanced levels of the IEOP at Iowa State University. Their TOEFL scores ranged from 420 to 617, with a mean of 498, and their Michigan English Language Placement Test (MELPT) scores from 65 to 92, with a mean of 76. Twenty-six of the participants were graduate students, nineteen were undergraduates, and five were nonacademic. The largest ethnic group, comprising nearly a third of the students, was Japanese, and the second largest group was Latin American. No academic participant in the study was more than two months away from full-time study.

Materials

The researchers selected three passages on scientific topics from a Scholastic Aptitude Test preparation manual (CEEB, 1990) based on their likely interest and language appropriateness for advanced ESL students in a pre-academic program. Though equal in length (approximately 375 words), the passages represented distinctively different expository types. Passage 1 described the life of Lloyd Hall, a pioneer in food chemistry; passage 2 contrasted the lives of ancient Pueblos and Navajos; and passage 3 speculated on possible causes of the Tunguska blast, which left a huge crater in Siberia.

Procedure

On the first day of the study participants answered written questions designed to provide information about their approaches to learning. They returned the second day to read one of the three passages under one of two possible conditions: in preparation to take a multiple-choice test (test condition) or to explain the passage to a fellow student (tutor condition). In all cases students were told that they could take as many notes as they wished in any form on the passage itself and/or on a separate sheet provided. They were allowed to use dictionaries of any kind and to ask for explanations from the two teacher-researchers in the room. Participants were given 20 minutes for each task, after which they completed a questionnaire that was designed to assess their strategies and the feelings they experienced while studying the passage, as well as its interest and difficulty for them. At this point, they either took a test or talked about their passage with a fellow student, according to their prior instructions. On the third day, students read a different passage under contrasting conditions. Remaining procedures were identical. The particular assignments to participants of passage, condition, and order in this balanced design enabled the researchers to determine the relationships between these factors and the notes taken.

Coding of data

After collecting notes on two passages from each student, each researcher independently examined a subset of these products and made lists of various notetaking strategies found, for example, underlining, highlighting, paraphrasing, listing, translating, and copying. Together they characterized each strategy as indicative of deep processing or surface processing (Marton & Saljo, 1976), depending on whether it appeared to represent an attempt to extract and/or relate main ideas (deep) or merely to focus on an isolated element of the text (surface). The observed categories of notes and their classification within the deep/surface processing framework are shown in Figure 1, and notes exemplifying deep and surface strategies are shown in Figures 2 and 3.

Surface Processing

1. Copying or transcribing word for word from original document

2. Focusing on vocabulary
   - -underlining individual words
   - -highlighting individual words
   - -writing down/listing discrete words
   - -defining individual words

Deep Processing

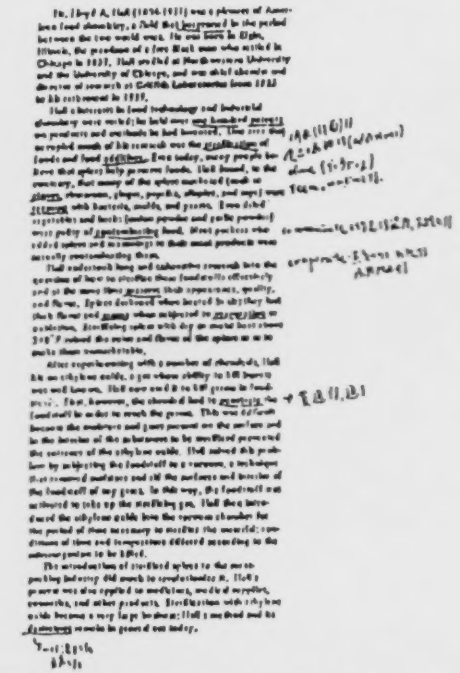
1. Organizing or diagramming information using graphic or pictorial devices including linking or separating lines, parentheses or brackets, space to separate concepts, arrow, bullets or footnotes

2. Paraphrasing from original article restating the text in another form or other words, either telegraphically (omitting function words) or in full sentences

3. Focusing on main ideas
   - -underlining main ideas
   - -highlighting main ideas
   - -writing down /listing main ideas

FIGURE 1. *Surface and deep processing strategies represented in notes*

A. Translating and vocabulary focus (underlining individual words)



B. Vocabulary focus (listing discrete words)

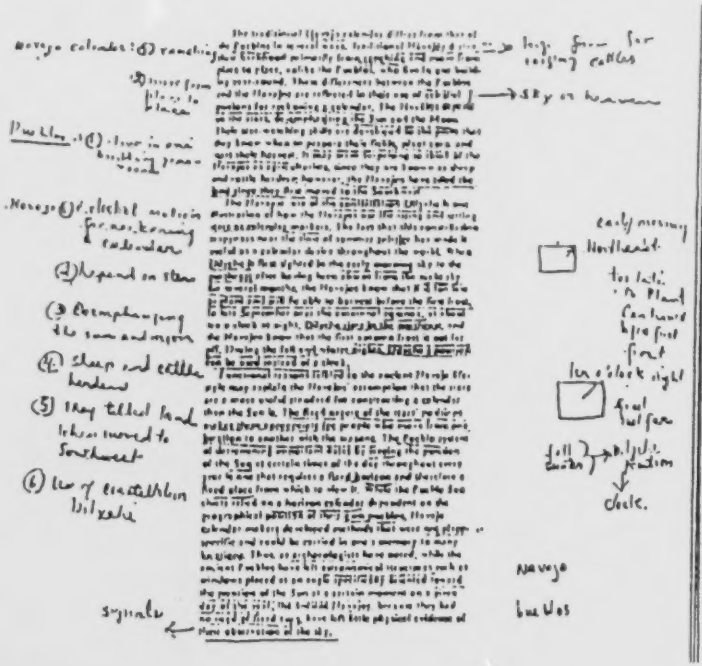
- he built over one hundred private
- sterilization
- additives
- cloves
- tannins • be full
- contaminating
- preserve • keep, conserve, reserve
- evaporation
- aroma
- penetrate
- derivatives

C. Translating and vocabulary focus (listing discrete words)

FIGURE 2. ***Examples of notetaking strategies: Surface processing***

974 BEST COPY AVAILABLE

A. Diagramming, paraphrasing, and main idea focus (underlining)



B. Diagramming, paraphrasing, and main idea focus

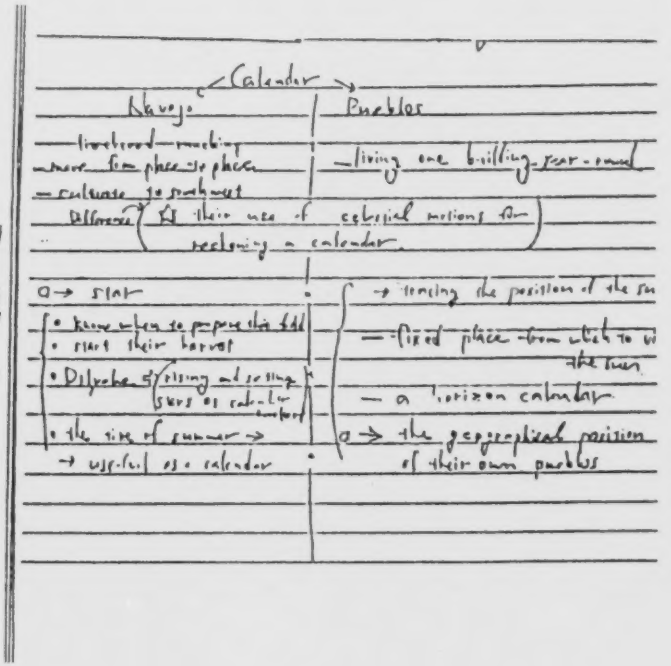

Calendar
Navajo
Pueblos
—herdsmen—working
—move from place to place
—cultivate to produce
—living one building-year-round
Difference (in their use of celestial motions for reckoning a calendar.)
a → start
• know when to prepare their fields
• start their harvest
• Dispute of rising and setting stars as calendar markers
• the time of summer →
→ useful as a calendar
→ tracing the positions of the sun
— fixed place from which to watch the sun
— a horizon calendar
a → the geographical position of their own pueblos

FIGURE 3. ***Examples of notetaking strategies: deep processing***

975

BEST COPY AVAILABLE

The researchers then independently analyzed the remaining notes using the protocol. In most cases, strategies were coded as being either present or not in a particular set of notes, with no attempt made to determine the frequency of occurrence. Only for translation and definitions, where units could be commonly agreed upon, were individual units tabulated and frequency examined. In translation, sets of characters written in another language and separated by white space were counted as single units, and for definitions the meaning of each word or phrase written in English was so regarded.

The overall reliability of coding, determined by the percentage of judgments on which the coders originally agreed, was 80%. Discrepancies were discussed until agreement was reached.

Analysis and Results

Correlations with other standard measures

As indicated in the literature on first-language reading reviewed earlier in this paper, deep processing has been associated with text comprehension and superior learning, and surface processing with lack of comprehension and learning. Depth of processing has also been specifically associated with the type of notes taken, with the notes of novices more likely to relate to the surface structure of the text and the notes of experts reflecting greater depth of processing. The question here is whether notes made by L2 learners while reading academic texts reflected the extent of their deep and surface processing, thus arguing for learner notes as a means of gaining insight into certain key learning processes. It was hypothesized that the notes of more proficient readers would reflect more deep processing strategies and fewer surface processing strategies than the notes of less proficient readers, and, conversely, that the notes of less proficient readers would reflect more surface and fewer deep strategies. Strong relationships between student performance on TOEFL and MELPT total and reading/vocabulary part scores and the type of notes produced would support these hypotheses.

To test the hypotheses, Vann and Schmidt used correlational procedures as described by Bohrnstedt and Knoke (1988; see also Borgatta & Bohrnstedt, 1981). The hypotheses were confirmed in part, with strong correlations in the expected directions for translation, copying, and diagramming. Simple correlation coefficients indicated a significant[12] negative correlation between translation and MELPT score (-.3418), MELPT vocabulary part score (-.3806), MELPT reading part score (-.4840), and TOEFL general score (-.4081), and in the expected direction, though not significant for TOEFL reading (-.2472), showing that translation as apparent in notes was associated with lower proficiency scores as expected. Likewise, copying was negatively correlated with vocabulary scores on the MELPT (-.3606) to a significant degree and negatively correlated with MELPT and TOEFL reading scores, though not significantly. Though vocabulary-focus correlations were in the expected direction, none was significant. For deep processing strategies, diagramming was positively correlated with both the structure (.4591) and reading (.5565) subparts of the TOEFL. Though not all categories

functioned equally effectively, overall the correlations between certain aspects of notetaking and proficiency suggest that information we can gain from student notes is in accord with information gained about reading skills from standardized tests. The fact that no statistically significant correlation for paraphrasing was found may have resulted from an excessively broad definition that characterized all notes not verbatim from the text as paraphrasing. True paraphrasing may, in fact, entail more sophisticated skills not captured with this definition.

Relationships with other learner variables

As another step in validating the classification of notes developed in this study, Vann and Schmidt examined whether certain categories of strategies might be connected with other key learner variables besides language and reading proficiency. Because prior studies have shown the importance of expertise (presumably acquired through experience) and suggested the role of language and culture in shaping learner strategies, these variables were examined. Learners with more academic experience (graduates) were expected to show more evidence of deep processing strategies than those with less academic experience (undergraduates). Students who came from cultures that emphasize copying and memorizing (Huang & Van Naerssen, 1987) and that might cause students to face greater challenges with vocabulary (in part because cognates with their L1 would be less common) were expected to show more evidence of surface strategies.

Undergraduates versus graduates. The hypothesis that graduate students would use more deep processing strategies than undergraduates was partially supported, as shown by the X^2 values in Table 3.[13] Although both graduates and undergraduates used the surface processing strategies of translating and copying verbatim from the texts, half of all undergraduates used these strategies, in comparison to fewer than 15% of the graduates who translated and 19% who copied. Undergraduates and graduates were, however, almost identical in their focus on vocabulary. For deep strategies, the pattern reversed, with more graduates than undergraduates using diagramming, paraphrasing, and main idea focus. Differences were in the direction predicted in all cases except for vocabulary focus, and were statistically significant for translation, copying, and main idea focus.

These differences provide partial empirical support for the validity of the classification of notes described here. They further suggest that certain key academic strategies may be acquired with academic experience, though it is also conceivable that users of deep processing strategies are more likely to become graduate students.

Asians versus Latin Americans. As expected, the notes of Asians revealed more surface processing strategies than did the notes of Latin Americans. As indicated in Table 4, most students did not translate, but those who did were primarily Asian, 37% of whom showed evidence of translation in their notes. Even more dramatically, 48% of all Asians showed evidence of copying directly from the text in their notes, while there was no evidence of copying among Latin Americans. Similarly, 85% of the Asians showed evidence of focusing on vocabulary in their notes, while only 38% of the

TABLE 3

Percentage (and number) of undergraduates and graduates taking notes reflecting surface and deep strategies (n=43)

Type of notes	Undergraduates (n=16)	Graduates (n=27)	X^2 (df)
Surface			
Translating	50% (8)	15% (4)	6.18 (1)
Copying	50% (8)	19% (5)	4.84 (1)
Vocabulary focus	69% (11)	70% (19)	.012 (1)
Deep			
Diagramming	63% (10)	81% (22)	1.90 (1)
Paraphrasing	81% (13)	93% (25)	1.17 (1)
Main idea focus	63% (10)	93% (25)	5.90 (1)

TABLE 4

Percentage (and number) of Asian and Latin American students taking notes reflecting surface and deep strategies (n=40)

Type of notes	Asian (n=27)	Latin American (n=13)	X^2 (df)
Surface			
Translating	37% (10)	8% (1)	3.79 (1)
Copying	48% (13)	0	9.27 (1)
Vocabulary focus	85% (23)	38% (5)	9.13 (1)
Deep			
Diagramming	70% (19)	92% (12)	2.42 (1)
Paraphrasing	85% (23)	92% (12)	0.41 (1)
Main idea focus	74% (20)	92% (12)	1.82 (1)

Latin Americans did. For deep processing strategies, although a greater percentage of Latin Americans used diagramming, paraphrasing, and main idea focus than did Asians, these differences were not significant. What is clear here is that Asians were more likely to reveal surface strategies in their notes: translating, copying, and focusing on vocabulary. Though we have speculated earlier as to possible reasons, when we keep in mind the relatively high proficiency levels of this group, it seems likely that prior educational training, as suggested by Huang & Van Naerssen (1987), and the extent to which there is shared vocabulary between the learner's first language and English both play a critical role.

Contextual Variables

There was no statistical support for the prediction that the passage, the student's assessment of the difficulty of the passage, the condition (test versus tutor), or the order in which the student read the passage influenced the extent of deep or surface processing. These findings would suggest that deep and surface processing are fairly strong traits not greatly influenced by at least some contextual modifications. They do not suggest that context has no effect, only that the kinds of contextual manipulations in this study did not result in perceivable differences in notetaking.

Discussion

This segment of the paper has focused on validity issues surrounding the development of a relatively untested tool for assessing academic reading strategies. The process of validation has so far consisted of attempting to anchor notetaking in relevant prior theory and connect it empirically with factors suggested by that theory. However, there are other features of reader notes that contribute to the validity of their use as a means of assessing learner strategies.

As a by-product of academic reading, notes are readily available to teachers and researchers for evaluation, thus enabling frequent in-class assessment. Comparing notes across individual assignments can provide information about the conditions under which a student chooses one strategy over another, and lead to a better understanding of the relationships between context and strategic decisions. However, the ease of collecting readers' notes is not merely convenient; it allows data to be collected without imposing constraints on the individual's reading such as insisting on oral reading or stopping the reader intermittently and asking for commentary unrelated to the passage. Reader notes offer two other advantages as well. Because notes, like other learner-initiated products, avoid self-report or retrospective data, they enable the analyst to tap strategies of which learners may not be conscious or which they may be unable or unwilling to report. The observer sees concrete evidence, not testimony. Finally, notes, somewhat like individual fingerprints, can be classified into types, but

remain unique, reflecting the individual response of a reader to a text in a way no standardized test can.

Although notes are limited to assessing academic reading strategies, it is difficult to overestimate the importance of academic reading for many of the world's users of English. Assuming that academic reading requires different processes from ordinary everyday reading (Carver, 1992), it would seem that specialized assessment tools are appropriate. As with other tools for strategy assessment, notetaking does not provide a direct view of cognitive processes; the researcher must infer these from the data the learner provides. Because the kinds of inferences we can make from each kind of assessment are different and limited, a variety of methods—think-alouds, interviews, product analysis—should be employed to obtain a comprehensive picture of learner strategies.

Given the limited population, texts, and contexts in the study, these findings must be interpreted tentatively. Further studies should be conducted to pursue key questions about the relationship between reader notes and reader behavior in various contexts. One obvious question currently being examined by Vann and Schmidt is the extent to which different kinds of notetaking behavior might correlate with differences in recalling texts. Other studies might pursue further how the learner's assessment of a task influences strategic behavior and presumably task performance. MacIntyre's (1994) framework for examining strategies in a psychological and social context may prove especially fruitful in helping future researchers formulate the complex relationships involved between individual traits and contextual variables.

Conclusion

The two studies discussed above suggest that common classroom products, responses to cloze tests and notes made by readers, can be generated under various experimental conditions and offer a practical and natural means for gaining access to certain types of learner processes. Following the approach suggested by Bialystok (1990a), the researchers analyzed the products within conceptual frameworks that permitted the illumination of underlying processes. The results were validated by a variety of means that demonstrated predicted relationships with learner and contextual variables. While products cannot reveal the entire breadth of learner processes, they can provide information not obtainable through other types of data collection and validate findings from other sources. Thus, the work reported here bears out the wisdom of Naiman et al.'s (1978) call for the use of both process *and* product data in understanding second language learning. We look forward to future studies demonstrating the validity of other uses of products to assess learner processes.

Notes

[1] This view of product differs from that in which product is "typically defined by global proficiency scores" (Larsen-Freeman & Long, 1991, p. 283). It should be noted that

such global measures usually require "selected" responses (in multiple-choice items), whereas the responses of interest here are "constructed" (Bachman, 1990).

[2] A number of SLA researchers, for example, Blum and Levenston (1978), Bialystok (1978), Frauenfelder and Porquier (1979), Faerch and Kasper (1983), and Seliger (1984)—all discussed in Bialystok (1990a)—and Rubin (1981) have distinguished between processes and strategies. Processes are typically described as "the mental steps taken to carry out a cognitive activity" (Bialystok, 1990a, p. 15), whereas strategies have been variously characterized as optional, conscious, and/or problem-oriented occurrences that come into play on a particular occasion. Bialystok (1990a) suggests that the criteria for differentiating processes from strategies in language production are "fragile" (p. 24) and that the two types of events may not be basically different.

[3] Tarone's 1977 taxonomy grew out of earlier attempts by Varadi (1973) and Tarone, Cohen, and Dumas (1976) to characterize communication strategies.

[4] There were no significant differences in the effectiveness of three of the four strategies tested (approximation, circumlocution, and language switch). The fourth strategy, word coinage, used sparingly by the learners who had produced the descriptions used in this study, was statistically more effective than the other three.

[5] To illustrate this confusion, Kellerman et al. (1990) refer to Varadi's (1973) example of a balloon being called an "airball" or a "ball with air in it," (Kellerman et al., 1990, p. 165) by a learner who does not know the correct word. Using typical taxonomic categories, the first would be classified as a word-coinage and the second, as paraphrase. However, the difference between the two attempts is merely one of superficial linguistic form.

[6] Here Kellerman et al.'s example is of a learner who, lacking the words *moon* and *knife*, substitutes "it's in the sky" and "you use it for cutting," respectively. Again, using carefully defined taxonomic categories, the first response would be classified as a description of a location, and the second as a description of function. However, this categorization "fails to capture the obvious point that the contents of the descriptions are only as different from each other as *the referents themselves*" (Kellerman et al., 1990, p. 166, italics in original).

[7] Subsequent revisions of the framework appear in Bialystok (1990a, 1990b, 1994). In these more recent versions, Bialystok regards the analyzed knowledge dimension as a *process*, with emphasis on the *analysis* of knowledge, in which mental representations, initially "loosely organized around meanings (knowledge of the world)," become "rearranged into explicit representations that are organized around formal structures" (Bialystok, 1994, p. 159). Thus, a learner's representations of knowledge at any given time are indicators of the degree of analysis that has taken place.

[8] The reports obtained in this study would fall into Cohen's (this issue) category of *self-observation*. Although participants were not explicitly told to think aloud in English, the context of the data collection strongly encouraged English, and it was exclusively used. If we repeated this study today, we would give participants a clear choice between using their native language or English.

[9] All learner names used in this report are pseudonyms.

[10] Poulisse, Bongaerts, and Kellerman (1987) discuss the usefulness of this kind of validation in the context of communication strategies.

[11] "Academic reading" is defined here as any kind of reading for academic purposes.

[12] Any reference to statistical significance implies alpha = .05.

[13] Of the initial sample of 50 participants, two students failed to complete the second reading and five students had "special student" status. Thus, strategies for a total of 43 undergraduates and graduates were analyzed here.

References

Abraham, R. G., & Chapelle, C. A. (1992). The meaning of cloze test scores: An item difficulty perspective. *Modern Language Journal, 76*(4), 468-479.

Abraham, R. G., & Vann, R. J. (1994, March). *Metacognition in the cloze performance of nine Lebanese ESL learners*. Paper presented at the Annual Conference of the American Association for Applied Linguistics, Baltimore.

Amer, A. A. (1994). The effect of knowledge-map and underlining training on the reading comprehension of scientific texts. *English for Specific Purposes, 13*(1), 35-45.

Anderson, T. H., & Armbruster, B. B. (1980). *Studying* (Technical Report No. 155). University of Illinois: Center for the Study of Reading.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly, 25*(4), 671-704.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*(1), 69-83.

Bialystok, E. (1983). Some factors in the selection and implementation of communication strategies. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 100-118). London: Longman.

Bialystok, E. (1990a). *Communication strategies: A psychological analysis of second-*

language use. Oxford: Basil Blackwell.

Bialystok, E. (1990b). The dangers of dichotomy: A reply to Hulstijn. *Applied Linguistics, 11*(1), 46-51.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16* (2), 157-168.

Bialystok, E,. & Frohlich, M. (1980). Oral communication strategies for lexical difficulties. *Interlanguage Studies Bulletin, 5*, 3-10. Cited in Bialystok (1990a).

Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition and human performance*, Vol. 1: *Theoretical perspectives* (pp. 207-252). Orlando, FL: Academic Press.

Block, E. L. (1992). See how they read: Comprehension monitoring of L1 and L2 readers. *TESOL Quarterly, 26*(2), 319 -343.

Blum, S., & Levenston, E. A. (1978). Universals of lexical simplification. *Language Learning, 28*(2), 399-415.

Blum-Kulka, S., & Levenston, E. A. (1983). Universals of lexical simplification. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 119-139). London: Longman.

Bohrnstedt, G. W., & Knoke, D. (1988). *Statistics for social data analysis* (2nd ed.). Itasca, IL: Peacock.

Bongaerts, T., Kellerman, E., & Bentlage, A. (1987). Perspective and proficiency in L2 referential communication. *Studies in Second Language Acquisition, 9*(2), 171-199.

Borgatta, E. F., & Bohrnstedt, G. W. (1981). Level of measurement: Once over again. *Sociological Methods and Research, 9*(2), 147-160.

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*(2), 145-153.

Brown, A. L., Bransford, J. D., Ferrera, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (4th ed.). Vol. 3: *Cognitive development* (pp. 77-166). New York: Wiley.

Brown, A. L., & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development, 49*(4), 1076-1088.

Brown, J. D. (1988). What makes a cloze item difficult? *University of Hawai'i Working Papers in ESL, 7*, 17-39.

Carver, R. (1992). Effect of prediction activities, prior knowledge, and text type upon amount comprehended: Using rauding theory to critique schema theory research. *Reading Research Quarterly, 27*(2), 165-174.

CEEB (1990). *10 SATS* (4th ed.). Princeton: College Entrance Examination Board.

Cohen, A. D. (1980). *Testing language ability in the classroom*. Rowley, MA: Newbury House.

Craik, F. I. M. (1977). Depth of processing in recall and recognition. In S. Dornic (Ed.), *Attention and performance* (pp. 679-697). Hillsdale, NJ: Lawrence Erlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Dechert, H. W. (1983). How a story is done in a second language. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 175-195). London: Longman.

Faerch, C., & Kasper, G. (1983). Plans and strategies in foreign language communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 20-60). London: Longman.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Erlbaum.

Frauenfelder U., & Porquier, R. (1979). Les voies d'apprentissage en langue etrangere. *Working Papers on Bilingualism, 17*, 37-64. Cited in Bialystok (1990a).

Galvan, J., & Campbell, R. (1979). An examination of the communication strategies of two children in the Culver City Spanish immersion program. In R. Andersen (Ed.), *The acquisition and use of Spanish and English as first and second languages*. Washington, DC: TESOL. Cited in Bialystok (1990a).

Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.

Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist, 24*(2), 143-158.

Haas, C. (1990). Composing in technological contexts: A study of note-making. *Written Communication, 7*(4), 512-547.

Haastrup, K., & Phillipson, R. (1983). Achievement strategies in learner/native speaker interaction. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 140-158). London: Longman.

Hidi, S., & Klaiman, R. (1983). Notetaking by experts and novices: An attempt to identify teachable strategies. *Curriculum Inquiry, 13*(4), 377-395.

Horwitz, E. K. (1987). Surveying student beliefs about language learning. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 119-129). Englewood Cliffs, NJ: Prentice Hall.

Huang, X-H., & Van Naerssen, M. (1987). Learning strategies for oral communication. *Applied Linguistics, 8*(3), 287-307.

Jamieson, J., & Chapelle, C. (1987). Working styles on computers as evidence of second-language learning strategies. *Language Learning, 37*(4), 523-544.

Jones, B. F., Pierce, J., & Hunter, B. (1988/89). Teaching students to construct graphic representations. *Educational Leadership, 46*(4), 20-25.

Kellerman, E., Ammerlaan, T., Bongaerts, T., & Poulisse, N. (1990). System and hierarchy in L2 compensatory strategies. In R. C. Scarcella, E. S. Andersen, & S. D. Krashen, *Developing communicative competence in a second language* (pp. 163-178). New York: Newbury House.

Kellerman, E., Bongaerts, T., & Poulisse, N. (1987). Strategy and system in L2 referential communication. In R. Ellis (Ed.), *The social context of second language acquisition*. Englewood Cliffs, NJ: Prentice-Hall. Cited in Bialystok (1990a).

Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research.* New York: Longman.

MacIntyre, P. D. (1994). Toward a social psychological model of strategy use. *Foreign Language Annals, 27*(2), 185-195.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology 46*(1), 4-11.

Marton, F., & Saljo, R. (1984). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning* (pp. 36-55). Edinburgh: Scottish Academic Press.

Naiman, N., Frohlich, M., Stern, H. H., & Todesco, A. (1978). *The good language learner.* Toronto: Ontario Institute for Studies in Education.

O'Malley, J . M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know.* New York: Newbury House.

Paribakht, T. (1985). Strategic competence and language proficiency. *Applied Linguistics, 6*(2), 132-146.

Poulisse, N. (1987). Problems and solutions in the classification of compensatory strategies. *Second Language Acquisition, 3*(2), 141-153.

Poulisse, N., Bongaerts, T., & Kellerman, E. (1987). The use of retrospective verbal reports in the analysis of compensatory strategies. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 213-229). Clevedon: Multilingual Matters.

Raupach, M. (1983). Analysis and evaluation of communication strategies. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 199-209). London: Longman.

Reid, J. M. (1987). The learning style preferences of ESL students. *TESOL Quarterly, 21*(1), 87-111.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 2*(2), 117-131.

Schmidt, H. H., & Vann, R. (1992, March). *Classroom format and student reading strategies: A case study.* Paper presented at the 26th Annual Convention & Exposition of Teachers of English to Speakers of Other Languages, Vancouver.

Seliger, H. W. (1984). Processing universals in second language acquisition. In F. R. Eckman, L. H. Bell, & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 36-47). Rowley, MA: Newbury House.

Snow, C., Cancini, H., Gonzalez, P., & Shriberg, E. (1989a). Giving formal definitions: An oral language correlate of school literacy. In D. Bloome (Ed.), *Classrooms and literacy* (pp. 233-249). Norwood, NJ: Ablex.

Snow, C. E., Cancini, H., Gonzalez, P., & Shriberg, E. (1989b). Second language learners' formal definitions: An oral language correlate of school literacy. In D. Bloome (Ed.), *Literacy in functional settings.* Norwood, NJ: Ablex. Cited in Bialystok (1990a).

Spivey, N. N. (1990). Transforming texts: Constructive processes in reading and wri-

ting. *Written Communication, 7*(2), 256-287.

Tang, G. M. (1992/1993). Teaching content knowledge and ESOL in multicultural classrooms. *TESOL Journal, 2*(2), 8-12.

Tarone, E. (1977). Conscious communication strategies in interlanguage. In H. D. Brown, C. A. Yorio, and R. C. Crymes (Eds.), *On TESOL '77* (pp. 194-203). Washington, D.C.: TESOL.

Tarone, E., Cohen, A., & Dumas, G. (1976). A closer look at some interlanguage terminology. *Working Papers in Bilingualism, 9*, 76-90. Cited in Tarone (1977).

Vann, R. J., & Abraham, R. G. (1990). Strategies of unsuccessful language learners. *TESOL Quarterly, 24*(2), 177-198.

Vann, R. J., & Schmidt, H. H. (1993, April). *Using student notes to assess reading strategies*. Paper presented at the 27th Annual Convention and Exposition of Teachers of English to Speakers of Other Languages, Atlanta.

Varadi, T. (1973). Strategies of target language learner communication: Message adjustment. Paper presented at the VIth Conference of the Romanian-English Linguistics Project, Timisoara. Cited in Kellerman et al. (1990).

Varadi, T. (1980). Strategies of target language learner communication: Message-adjustment. *International Review of Applied Linguistics, 18*(1), 59-72.

Wagner, J. (1983). Dann du tagen eineeeee—weisse Platte—an analysis of interlanguage communication in instructions. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 159-174). London: Longman.

Wenden, A. (1987). Metacognition: An expanded view on the cognitive abilities of L2 learners. *Language Learning, 37*(4), 573-597.

Yussen, S. R. (1985). The role of metacognition in contemporary theories of cognitive development. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance,* Vol. 1: *Theoretical perspectives* (pp. 253-283). Orlando, FL: Academic Press.

Acknowledgment

We wish to thank the anonymous reviewers and Lidia Woytak for helpful suggestions on an earlier version of this article.

Authors

ROBERTA G. ABRAHAM, Professor Emerita, Department of English, Iowa State University, 203 Ross Hall, Ames, IA 50011-1201. Specialization: individual differences in second language learning.

ROBERTA J. VANN, Professor, Department of English, Iowa State University, 203 Ross Hall, Ames, IA 50011-1201. Specialization: methods of teaching English as a second language.

General Information

Calendar of Events*

1996

17-18 October, *Minnesota Council on the Teaching of Foreign Languages,* location to be announced. Information Millie Mellgren, 3055 Rosewood Lane, Plymouth, MN 55441.

18-20 October, *New York State Association of Foreign Language Teachers,* Kiamesha Lake. Information Helene Combopiano, 71 Grand Boulevard, Binghamton, NY 13905.

18-20 October, *Texas Foreign Language Association,* Houston. Information TFLA, 1320 Modiste Drive, Houston, TX 7705; (713)468-4959.

23-25 October, *Foreign Association of North Dakota,* Bismarck. Information Herbert Boswau, P.O. Box 8198, Grand Forks, ND 58202-8198.

24-26 October, *Foreign Language Association of North Carolina,* Greensboro. Information Wayne Figart, P.O. Box 739, Wilmington, NC 28402-0739; (910)763-5733.

24-27 October, *Second Language Acquisition and Teaching,* Tucson. Information SLAT, Modern Languages 347, University of Arizona, Tucson, AZ 85721; Email [SLRF96@ccit.arizona.edu].

1-2 November, *National Association of Self-Instructional Language Programs,* Washington. Information NASILP, Critical Languages, 022-38, Temple University, Philadelphia, PA 19122.

22-24 November, *American Council on the Teaching of Foreign Languages with American Association of Teachers of German,* Philadelphia. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6810; Fax (914)963-1275.

27-30 December, *Modern Language Association of America,* Washington, D.C. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212)477-9863.

1997

9-11 January, *Chronos Conference,* Brussels. Information Svetlana Vogeleer, Institut Libre Marie Haps, rue d'Arlon 3, 1040 Brussels, Belgium; Fax 32-2-511-98-37.

6-8 March, *Southern Conference on Language Teaching with South Carolina Foreign Language,* Myrtle Beach; Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912)333-7358, Fax (912)333-7398, Email [lbradley@grits.valdosta.peachnet.edu].

8-11 March, *American Association of Applied Linguistics,* Orlando. Information AAAL, 7630 West 145th Street, Suite 202, Apple Valley, MN 55124; (612)953-0805, Fax (612)891-1800, Email [howe@mr.net].

* Courtesy of *The Modern Language Journal* (University of Wisconsin Press)

11-15 March, *Teachers of English to Speakers of Other Languages,* Orlando. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703)836-0774, Fax (703)836-7864, Email [tesol@tesol.edu].

3-5 April, *North American Association of Christian Foreign Language and Literature Faculty,* Grand Rapids. Information Barbara Carvill, Calvin College, Grand Rapids, MI 49546; (616)957-6365, Email [carv@calvin.edu].

3-6 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, St. Michael's College, 29 Ethan Allen Avenue, Colchester, VT 05446.

3-6 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717)245-1977, Fax (717)245-1976, Email [neconf@dickinson.edu].

10-12 April, *Pacific Northwest Council for Languages,* Eugene. Information PNCFL, Foreign Languages and Literatures, Oregon State University, 210 Kidder Hall, Corvallis, OR 97331-4603; Fax (541)737-3563, Email [verzascr@cla.orst.edu].

10-13 April, *Central States Conference on the Teaching of Foreign Languages,* Columbus. Information CSCTFL, Madison Area Technical College, 3550 Anderson Avenue, Madison, WI 53704; (608)246-6573, Fax (608)246-6880.

10-13 April, *Central States Conference on the Teaching of Foreign Languages,* Columbus. Information CSCTFL, Rosalie Cheatham, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; (501)569-8159, Fax (501)569-3220, Email [rmcheatham@ualr.edu].

16-18 April, *Russia's Golden Age,* Columbus. Information George Kalbouss, Slavic & E. European Languages & Literatures, 232 Cunz Hall, 1841 Millikin Road, Ohio State University, Columbus, OH 43210; (614)292-6733, Fax (614)292-3107.

18-20 April, *Midwest Slavic Conference,* Columbus. Information George Kalbouss, Slavic & E. European Languages & Literatures, 232 Cunz Hall, 1841 Millikin Road, Ohio State University, Columbus, OH 43210; (614)292-6733, Fax (614)292-3107.

25-26 April, *Classical Association of the Atlantic States,* Wilmington. Information R. Leon Fitts, Department of Classical Studies, PO Box 1773, Dickinson College, Carlisle, PA 17013-2896.

8-10 May, *Non-Slavic Languages,* Chicago. Information NSL-10, Department of Slavic Languages & Literatures, University of Chicago, 1130 E. 59th St., Chicago, IL 60637; (312)702-8033, Fax (312)702-9861, Email [hia5@midway.unchicago.edu].

4-7 July, *Society for the History of Authorship, Reading and Publishing,* Cambridge. Information SHARP, 51 Sherlock Close, Cambridge CB3 0HP, United Kingdom.

5-7 August, *Problems of Teaching Modern Hebrew,* Jerusalem. Information Ben-Zion Fischler, Council on the Teaching of Hebrew, P.O.B. 7413, Jerusalem 91073, Israel.

12-16 August, *Foreign Language Education and Technology,* Victoria. Information (604)721-8294, Fax (604)721-8778, Email [fleatIII@call.uvic.ca], WWW [http://ikra.call.uvic.ca/langcen/fleat/fleat1.html].

4 September, *Association of Literary Semantics,* Freiburg. Information Monika Fluderik, English Department, University of Freiburg, D-79085, Germany.

7-11 October, *Foreign Language Association of North Carolina,* Winston-Salem. Information Wayne Figart, P.O. Box 739, Wilmington, NC 28402-0739; (910)763-5733.

7-11 October, *Foreign Language Association of North Carolina,* location to be announced. Information Mary Lynn Redmond, 6 Sun Oak Court, Greensboro, NC 27410; Fax (910)759-4591, Email [redmond@wfu.edu].

21-23 November, *American Council on the Teaching of Foreign Languages,* Nashville. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914)963-8830, Fax (914)963-1275.

27-30 December, *Modern Language Association of America,* Toronto. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212)477-9863.

27-30 December, *American Association of Teachers of Slavic & E. European Languages,* Toronto. Information AATSEEL, 1981 N. Evelyn Ave., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com].

1998

26-28 February, *Southern Conference on Language Teaching with Foreign Language Association of Georgia,* Savannah. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912)333-7358, Fax (912)333-7389, Email [lbradley@grits.valdosta.peachnet.edu].

17-21 March, *Teachers of English to Speakers of Other Languages,* Seattle. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2705; (703)518-2521, Fax (703)836-7864, Email [jennifer@tesol.edu].

26-29 March, *Central States Conference on the Teaching of Foreign Languages,* Milwaukee. Information CSCTFL, Rosalie Cheatham, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; (501)569-8159, Fax (501)569-3220, Email [rmcheatham@ualr.edu].

16-19 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717)245-1977, Fax (717)245-1976, Email [neconf@dickinson.edu].

24-26 April, *Pacific Northwest Council for Languages,* Boise. Information PNCFL, Foreign Languages and Literatures, Oregon State University, 210 Kidder Hall, Corvallis, OR 97331-4603; Fax (541)737-3563, Email [verzascr@cla.orst.edu].

September dates to be announced, *International Congress of Slavists,* Cracow. Information Robert Rothstein, Department of Slavic Languages, Herter Hall, University of Massachusetts, Amherst, MA 01003-3940.

27-31 October, *Foreign Language Association of North Carolina,* Winston-Salem. Information Wayne Figart, P.O. Box 739, Wilmington, NC 28402-0739; (910)763-5733.

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (*ALL*) is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes research and review articles, research reports as well as reviews on adult language learning for functional purposes. The editor encourages the submission of research and review manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Research Article

Divide your manuscript into the following sections:

- Abstract
  - Introduction
    - Method
      - Results
        - Discussion
          - Conclusion
            - Appendices
              - References
                - Author

Introduction

In a few paragraphs state the purpose of the study and relate it to the hypothesis and the experimental design. Point out the theoretical implications of the study and relate them to previous work in the area.

Next, under the subsection *Literature Review,* discuss work that had a direct impact on your study. Cite only research pertinent to a specific issue and avoid references with only tangential or general significance. Emphasize pertinent findings and relevant methodological issues. Provide the logical continuity between previous and present work. Whenever appropriate, treat controversial issues fairly. You may state that certain studies support one conclusion and others challenge or contradict it.

Method

Describe how you conducted the study. Give a brief synopsis of the method. Next develop the subsections pertaining to the *participants,* the *materials,* and the *procedure.*

Participants. Identify the number and type of participants. Specify how they were selected and how many participated in each experiment.

Identify the number of experiment dropouts and the reasons they did not continue.

Provide major demographic characteristic such as age, sex, geographic location, and institutional affiliation. Indicate that the treatment of participants was in accordance with the ethical standard of the APA (Principle 9).

Apparatus. Describe briefly the materials used and their function in the experiment.

Procedure. Describe each step in the conduct of the research. Include the instructions to the participants, the formation of the groups, and the specific experimental manipulations.

Results

First state the results. Next describe them in sufficient detail to justify the findings. Mention all relevant results, including those that run counter to the hypothesis.

Tables and figures. Prepare tables to present exact values. Use tables sparingly. Sometimes you can present data more efficiently in a few sentences than in a table. Avoid developing tables for information already presented in other places. Prepare figures to illustrate key interactions, major interdependencies, and general comparisons. Indicate to the reader what to look for in tables and figures.

Discussion

Express your support or nonsupport for the original hypothesis. Next examine, interpret, and qualify the results and draw inferences from them. Do not repeat old statements: Create new statements that further contribute to your position and to readers understanding of it.

Conclusion

Succinctly describe the contribution of the study to the field. State how it has helped to resolve the original problem. Identify conclusions and theoretical implications that can be drawn from your study.

Review Article

It should describe, discuss, and evaluate several publications that fall into a topical category in foreign language education. The relative significance of the publications in the context of teaching realms should be pointed out. A review article should be 15 to 20 double-spaced pages.

Research Report

It should present and discuss data obtained in a research project in the area of foreign language education. A research report should be 5 to 10 double-spaced pages.

Review

Reviews of textbooks, scholarly works on foreign language education, dictionaries, tests, computer software, video tapes, and other non-print materials will be considered for publication. Both positive and negative aspects of the work(s) being considered should be pointed out. The review should give a clear but brief statement of the work's content and a critical assessment of its contribution to the profession. Quotations should be kept short. Reviews that are merely descriptive will not be accepted for publication. The length of the manuscript should be three to five double-spaced pages.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-*ALL*
ATTN: Editor (Dr. L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not being considered for publication elsewhere.

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784. Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run from 30 to 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. An abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

A list of references should be submitted on a separate page of the manuscript with the centered heading: *References*. The entries should be arranged alphabetically by surname of authors. The sample list of references below illustrates the format for bibliographic entries:

Dulay, H. & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 16* (1), 93-95.

Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

Reference citations in the text of the manuscript should include the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that are being quoted originally appeared, e.g., (Jones, 1982, pp. 235-238). All works cited in the manuscripts must appear in the list of references, and conversely, all works included in the list of references must be cited in the manuscript.

Notes should be used for substantive information only, and they should be numbered serially throughout the manuscript. They all should be listed on a separate page entitled *Notes*.

Specifications for Floppy Disks

Where feasible, manuscripts are preferred on Windows-based software. Manuscripts produced on one of the DOS-based or Macintosh systems should be formatted as an ASQII file at double density, if possible.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____
2. Disk is formatted as: double__ high density__
3. Remarks:_____

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

994

GPO 583-777/79045

PIN: 074 768 - 000

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

BULK RATE
U.S. POSTAGE
PAID
MONTEREY, CA
PERMIT NO. 42

996

ERIC Clearinghouse on Langs. & Lings.
1118 22nd Street, N.W.
Washington, DC 20037

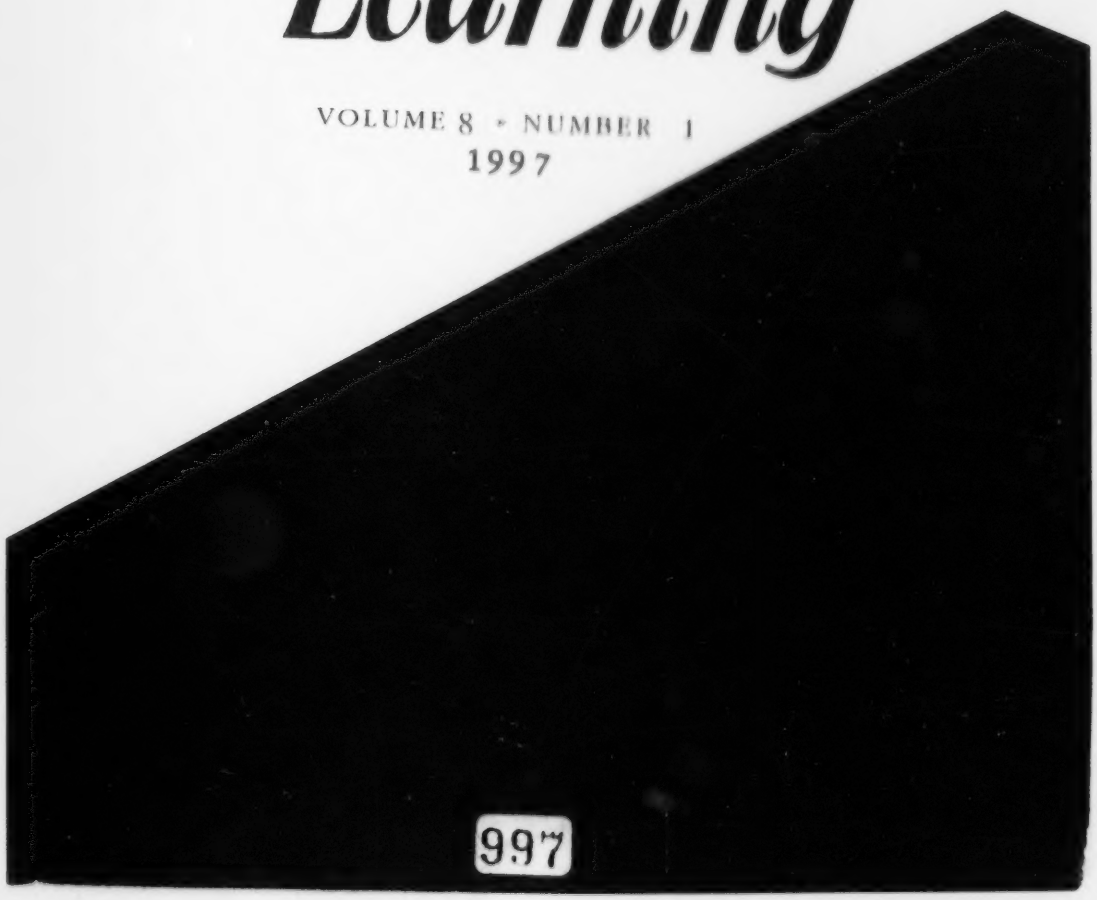
Professional Bulletin 65-96-1&2
United States Army
Approved for public release.
Distribution is unlimited.

996

BEST COPY AVAILABLE

Applied Language Learning

VOLUME 8 • NUMBER 1
1997



997

Applied Language Learning

1997 Volume 8 Number 1

Editor

Lidia Woytak

Defense Language Institute Foreign Language Center
Presidio of Monterey

998

Applied Language Learning

PB 65-97-1

The mission of Professional Bulletin 65-97, *Applied Language Learning (US* ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semiannually by the Defense Language Institute Foreign Language Center and Presidio of Monterey, presents professional information: The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

Officials:

DENNIS J. REIMER
General, United States Army
Chief of Staff

JOEL B. HUDSON
Administrative Assistant to the
Secretary of the Army
03598

DANIEL D. DEVLIN
Colonel, United States Army
Commander and Commandant, DLIFLC

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the authors.

Availability

To access *Applied Language Learning* on the Internet type

http://lingnet.army.mil

Additionally, you may obtain the journal on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster

Send change-of-address information to

Applied Language Learning
Defense Language Institute Foreign Language Center
ATFL: APAJ
Presidio of Monterey, CA 93944-5006

United Parcel Customers

Identify location of *Applied Language Learning*
Tin Barn
Bldg. 518, Room 6

Readers

Contact Editor, Dr. Woytak, by:

E-mail: woytakl@pom-emh1.army.mil
Telephone: 408 242-5638
DSN: 878-5638
Fax: 408 242-5850

Wordprocessing and formatting by **Carol Liles**
Cover design by **Barney Inada**

1000

Applied Language Learning
PB-65-97-1
1997 Volume 8 Number 1

1 The Effects of Lexical and Grammatical Cues on Processing Past Temporal Reference in Second Language Input
James F. Lee, Teresa Cadierno, William R. Glass, and *Bill VanPatten*

25 A Link Between Reading Proficiency and Native-Like Use of Pausing in Speaking
Ruth Johnson and *Rita Moore*

43 A Gender-Related Analysis of Strategies Used to Process Written Input in the Native Language and a Foreign Language
Dolly Jesuita Young and *Rebecca Oxford*

75 Pedagogical Intervention and Pragmatic Competence Development
Virginia LoCastro

Research Report

111 Free Voluntary Reading as a Predictor of TOEFL Scores
Rebecca Constantino, Sy-Ying Lee, Kyung-Sook Cho, and *Stephen D. Krashen*

Reviews

119 The Current State of Interlanguage: Studies in Honor of William Rutherford
J. S. Hedgcock

128 Taylor: Writing and Literacy in Chinese, Korean, and Japanese
J. Kuo and S. Akutsu

132 Reeder, Shapiro, Watson, and Goelman: Literate Apprenticeships
L. Plakans

General Information

135 Calendar of Events

141 Information for Contributors

1001

Applied Language Learning
1997, Vol. 8, No. 1, pp. 1-23

The Effects of Lexical and Grammatical Cues on Processing Past Temporal Reference in Second Language Input

James F. Lee
The Indiana University

Teresa Cadierno
Oden University

William R. Glass
McGraw Hill

Bill VanPatten
The University of Illinois at Urbana-Champaign

In this article we examine the hypothesis that learners' focal attention when processing input for semantic information is on lexical items as opposed to grammatical items (e.g., morphology). We present the findings of an investigation in which university students of Spanish from three levels of instruction make tense assignments under two textual conditions: adverbs plus verb inflection and verb inflection only. Processing is measured via two tasks: reconstruction of propositional content and recognition of the tense in which a verb appeared in the passage. The results for both tasks show consistent level differences. Lexical cues significantly improved the reconstruction of the propositional content of the passage learners listened to. These lexical cues did not, however, significantly improve tense recognition.

1002

© 1997, James F. Lee, Teresa Cadierno, William R. Glass, and Bill VanPatten

In an extensive review of the literature on effects of input modifications on comprehension, Long and Ross (1993) question whether what is good for enhancing comprehension (that is, linguistically simplifying input) is also good for language learning (that is, building a second language linguistic system), especially if the modifications to the input remove the very forms that need to be acquired. They are not alone in questioning the role of comprehensible input in second language (L2) learning (Barasch & James, 1994; Gass & Madden, 1985; Leow, 1995; Loschky, 1994). Following on Krashen's (1982) declaration of the insufficiency of comprehensible input alone to move acquirers along to later stages of development, researchers have sought to uncover what else is necessary for language acquisition to take place (Ellis, 1985; Gass, 1988; Larsen-Freeman & Long, 1991; Lightbown & Spada, 1993; Swain, 1985; VanPatten, 1992, 1996). Researchers began to problematize the relationship between comprehending the propositional content of input and utilizing the grammatical forms in the input to build linguistic systems. The clearest statement about these two phenomena is offered in Lee and VanPatten (1995).

> Although related, comprehension and input processing are not the same phenomenon. We define comprehension as making or creating meaning from the informational content in the input [for the purpose of interpreting a message]. We define input processing as making form-meaning connections from the linguistic data in the input for the purposes of constructing a linguistic system. (p. 96)

Input processing has recently become the focus of research in second language acquisition (SLA) circles and has been concerned with the psycholinguistic aspects of input. That is, it has set out to examine issues such as how learners process L2 input data, that is, how learners make form-meaning connections during comprehension, what types of strategies are used by these learners when attending to linguistic data in the input, and how and what part of that input becomes intake. Input processing research examines what aspects of language learners attend to, notice, detect, and process so that they are somehow registered in working memory. From these grammatical features called intake, learners may engage in further processing in order to build and restructure their linguistic systems (Tomlin & Villa, 1994; VanPatten, 1996).

While Lee and VanPatten (1995) postulate that logically there must be some overlap between comprehension and input processing, the

1003

exact nature of the relationship between the two has yet to be explored. Typically, the two sets of processes are not brought into the same research design. Two examples should suffice to make this point. Lee (1987) examined the effects on comprehension of known versus unknown morphological forms in written input. He found that when the informational content of the unknown forms was extracted from the input; learners comprehended the meaning. Leow (1993) examined the effects of simplifying input on the recognition of linguistic forms from the input. He found that simplification did not affect recognition of forms after reading two passages. Whereas Lee can offer no evidence that the forms themselves were processed for use in building a linguistic system, Leow can offer no evidence that the recognized forms were comprehended or used to interpret the message.

In this article we examine both comprehension and input processing in the same design. We investigate the effects of lexical cues and grammatical cues on comprehension and input processing. We expose learners in one group to input characterized by morphological markings plus lexical cues and another group to the same input minus the lexical cues. We measure the effects of the exposure conditions via two tasks: reconstruction of propositional content and recognition that a piece of information from the text referred to the past.

Background

The Role of Attention

The role of attention in processing incoming linguistic data has been examined in both L1 and L2 research. In L1 acquisition, Slobin (1985) argued that the only linguistic material that can be acquired is the one that has attracted the child's attention, that is the one that has been noticed and held in memory. In other words, children must attend to language data for the data to become part of their linguistic system. With respect to SLA, Schmidt (1990, 1994) has concluded that learning without attention to what is to be learned is impossible and that, consequently, adult language learners must pay attention to form in order for that form to become intake. "You can't learn a foreign language (or anything else for that matter) through subliminal perception." (1990, p. 142)

In a recent article, Tomlin and Villa (1994) concluded that attention involves at least three separable components: alertness, which represents an overall and general readiness to deal with incoming stimuli; orientation, which is the specific aligning of attention on a stimulus,

that is, actually attending to that stimulus; and detection, which is the process that selects or engages a particular bit of information. Detection is, then, a subprocess of attention that, according to Schmidt (1994), is close or identical to what is known as focal attention.

Three key ideas about detection are relevant for the present discussion. One of the key ideas proposed by Tomlin and Villa states that "detected information is available for other cognitive processing." (1994, p. 192). That is, particular data must be detected before other cognitive processing can occur. According to the authors, detection is of great interest to SLA because it constitutes the process by which particular exemplars are registered in memory and therefore can be made accessible to key processes for learning such as hypothesis formation and testing. Detected information then would be available for reconstructing propositional content as well as for other subsequent memory-based tasks such as identifying the temporal reference of an event described in a passage.

Two other key ideas about detection state that "information detected causes great interference with the processing of other information" and "information detected (cognitive registration) exhausts more attentional resources than orientation of attention" (Tomlin & Villa, 1994, p. 192). In other words, detecting one bit of information consumes attentional resources, and therefore may interfere with the detection of other bits of information. Both ideas refer to the limited-resource metaphor of attention, according to which the human attentional system is conceived as a limited mental resource or capacity,

In the field of applied linguistics, the conception of L2 learners as limited-capacity information processors is not novel. As McLaughlin, Rossman and McLeod (1983) have argued, humans are limited in what they can attend to at a given point in time, and in what they can handle on the basis of their current knowledge and expectations. In other words, attention is selectively focused when humans process incoming stimuli (James, 1890).

What Do L2 Learners Attend to in the Input?

Given this limited capacity that all humans have, one must wonder how L2 learners "choose" what to process from the input to which they are exposed. Since the input processing capacity of learners is limited, only certain features receive attention at any given time during the processing of a sentence. But which features do learners attend to and, is attention directed in a principled way?

1005

VanPatten (1985, 1992, 1996) has addressed this question and has proposed a set of principles concerning what learners attend to in the input. The first principle is that learners process input for meaning before they process it for form. That is, learners are driven to look for the message in the input (propositional content) before looking for how that message is grammatically encoded. As VanPatten notes, this principle is consistent with the observations of other researchers in both L1 and L2 acquisition (Peters, 1985; Sharwood Smith, 1985, 1993) that state that both L1 children acquirers' and adult L2 learners' attention is directed toward meaning when processing input. In other words, when the learners' aim is to extract meaning from the input they will attend to those aspects of the input that aid them in this endeavor.

This principle has led VanPatten to postulate a second one stating that learners process content words in the input before anything else. More explicitly, since learners intuitively know that content words are the building blocks of meaning, their attentional resources will be directed toward the detection of these words for constructing the meaning of sentences. This principle is borne out in both L1 and L2 acquisition research in that learners in input-rich environments tend to pick out and start using single words and/or whole unanalyzed chunks of language, which they treat as content words. Mangubhai (1991) reports on subjects' introspection data gathered while they were learning Hindi via Total Physical Response. His learners report that they rely heavily on lexical cues and 'chunking' (parsing whole phrases and expressions as one lexical item) to comprehend the propositional content while ignoring grammatical cues.

Orienting Attention

Additional evidence for the primacy of learners' attentional focus on lexical content words when processing input for meaning comes from experimental studies in which subjects' attention is oriented toward detecting linguistic features in the input while simultaneously processing for propositional content. VanPatten (1990) showed that learners can be directed to attend to key lexical items without loss of propositional content when processing input but cannot be directed to attend to grammatical markers (articles, person, number, inflections on the verb) without loss of propositional content. Bransdorfer (1991) demonstrated the same phenomenon directing learners' attention toward lexical items versus the Spanish copular verb *estar*. In his study, learners were directed to attend to the lexical item *exámenes* (exams) without loss of propositional content but lost propositional content when

directed to attend to the copular verb and the propositional content at the same time. Hulstijn (1989) also provided learners directions orienting their attention to meaning and form. Learners were instructed to copy down interrogative sentences projected onto a screen. The form group copied the sentence and was given each sentence in eight fragments from which they were to reconstruct the original sentence. The meaning group copied the sentence and was directed to give their opinion of the question posed in the sentence. The *form + meaning* group was told to pay attention to both structure and meaning but they had no secondary task to perform. The dependent variable in this study was cued recall of the target sentences. The results demonstrated that all groups copied down the sentences equally well indicating equal ability across groups to register the input. The form group recalled the structure of the target sentences better than the other two groups but recalled the content of them worse than the others. Directing attention to form (via the scrambled sentence fragments) resulted in a loss of propositional content.

A third principle VanPatten proposes is that learners prefer processing lexical items to grammatical items (e.g., morphology) for semantic information. This hypothesis receives support from two kinds of evidence in L2 research. The first piece of evidence comes from research in the acquisition of tense. Tense can be encoded in lexical items and expressions (e.g., yesterday, two years ago, tomorrow) but can also be encoded grammatically (e.g., walked vs. will walk). What research shows (Bardovi-Harlig, 1992; Klein, 1986, 1994) is that learners typically mark time early in the acquisition of verb morphology through lexical items (yesterday, last week) and only later start adding past tense verb markings. The second piece of evidence comes from Musumeci's (1989) study of Italian, French, and Spanish learners who were asked to assign tense to input sentences. Learners indicated whether what they heard (and saw) referred to the present, past, or future. The input was delivered orally under four conditions. The first three conditions called for verb morphology accompanied by: (1) adverbials of time; (2) typical teacher physical gestures; 3) both adverbials of time and physical gestures. The fourth condition called for verb morphology as the only source of information about the tense of the sentence. Musumeci's results were quite clear: The significant factor determining correct tense assignment was the presence or absence of temporal adverbials in the input sentences. Thus, her results provide evidence for the precedence of L2 learners to lexical items when attending to meaning in the input.

1007

The Present Study

In this article we continue the investigation of how learners process tense cues in the input using design features from the previous research materials as well as design features specific to this study. The result of manipulating the input data is that they often do not resemble "real world" language. Learners, both in and out of the classroom, are not exposed to decontextualized isolated sentences (cf. Hulstijn, 1989; Musumeci, 1989) but rather to sentences in context (cf. Bransdorfer, 1991; VanPatten, 1990). In our research design, we use a narrative of the type that learners might hear in the classroom as part of, for example, a listening comprehension activity. Our research paradigm also differs from previous research in another way. Rather than asking learners to indicate whether a sentence was coded for past, present, or future (Musumeci, 1989) or having them recall propositional content (Bransdorfer, 1991; VanPatten, 1990), we require both. We first ask learners to reconstruct the propositional content in order to determine what they actually have processed and stored as the tense of particular events in the narrative and then ask them to indicate the tense of particular events mentioned in the narrative. Previous research on discourse-level processing has not manipulated the characteristics of the input (Bransdorfer, 1991; Hulstijn, 1989; VanPatten, 1990), but research on sentence processing has (Musumeci, 1989). We decided, then, to manipulate the narrative to create two exposure conditions based on the presence or absence of lexical cues to tense. Contrary to previous research on discourse-level processing, we do not offer different orienting directions to the learners. They all listen as they naturally would to the meaning of the text (the caveat being how natural is any experiment). Learners from three levels of language experience were used in order to determine whether a developmental pattern would emerge.

Method

Subjects

University students enrolled in three different semester-level Spanish classes participated in the study. The levels represented were first, third, and fifth semester. Students enrolled in first and third semester levels were taught according to a communicative methodology modeled on the Natural Approach in the first semester and an expansion of Natural Approach in the third. In both these semesters all learners were taught

1008

using a common syllabus and were tested using common exams. First and third semester courses met four days a week. Students enrolled in the fifth semester level, on the other hand, were enrolled in an advanced grammar course where the formal learning of language structures was emphasized. This course met three days a week. Typically, students enrolled in the fifth semester course have completed four years of high school Spanish and have attained a high enough placement score to be placed in the fifth semester. A questionnaire was used to help delineate the subject pool. Specifically, students who reported that they used Spanish outside the classroom or had difficulty in hearing were eliminated from inclusion in the experiment. The number of students on whom data was analyzed was 102.

Materials

Each learner received a packet containing a consent form, background questionnaire, blank sheets to write the recalls (reconstructions of propositional content), and the tense identification task. Learners listened to two narratives. Although they were not told so, the first narrative served merely as a warm up to acquaint learners with the voice on the tape and the experimental procedure. Data were gathered and analyzed only on the basis of the second narrative, a two-minute passage about the singer and actor, Rubén Blades.

Two different conditions were used for the passage. Under Condition 1 the narrative contained targeted sentences in which both cues to past temporal reference were present: lexical (i.e., adverbs) and grammatical (i.e., verb inflections). The seven targeted verbs were in the preterite tense; all other verbs in the passage were either in the present tense or in the periphrastic future. A variety of adverbs was used, ranging from concrete reference (*ayer,* " yesterday") to a more remote abstract time frame (*durante los años sesenta* ,"during the 1960's"). Under Condition 2 the narrative was identical in content but all temporal adverbs were deleted. In order to process past temporal references, learners had to attend to the grammatical markers (i.e., verb inflections). (See Appendix A for the text of the two narratives.) The text with the lexical cues may seem a more natural sounding sample of connect discourse than the narrative without adverbials, although both versions are constructed of perfectly grammatical sentences. We decided, therefore, to examine the reconstructions of propositional content only for the presence of the seven targeted items and not to assess global comprehension. How well the total passage was comprehended is not the issue of concern; how *grammatical + lexical cues* versus *grammatical cues only* affect processing is.

1069

One-half of the learners from each level listened to the narrative under Condition 1 while the other half listened to the narrative under Condition 2. In order to facilitate comprehension, the learners were informed that the narrative contained information on the singer and actor Rubén Blades. They were not told the type of information nor were they given any details of the narrative prior to listening. All information and instructions given to the learners were in English, their native language.

The experimentation was carried out in the learners' classrooms during their regular class sessions. In all classes one of the experimenters was present. The passage was recorded by a near-native speaker and played to the learners on a stereo-cassette recorder. It should be noted that while the speech style of the speaker attempted to imitate that of a typical radio talk-show personality giving a report similar to the passage used, in no instance was any targeted item given suprasegmental emphasis to enhance acoustical salience. Before listening to the passage, the learners were instructed to listen to the content of the narrative and were told that their comprehension of what was said would be assessed afterwards.

Assessment

Free Recall Reconstruction of Content. After hearing the narrative, learners immediately performed a free recall by writing down in English, what they remembered about the passage. They were encouraged to write down any information that they could remember no matter how general or specific the information. Quantity was stressed as opposed to organizational quality. Full sentences were not required.

Tense Identification Test. Following the free recall, learners answered an eleven-item multiple-choice tense identification test. Each item required the learners to indicate whether specific information from the narrative was presented in the present, past, or future tense. An example follows (the entire test appears in Appendix B). The infinitive form of the verb was used because it is the citation form (that is, the lexical entry in a dictionary), and, being free of tense markings, the infinitive conveys only the lexical meaning of the verb.

VOLVER a la universidad

a. present
b. past
c. future

A potential confounding effect on learners' performance is the discourse frame of the passage. The grammatical subject of all seven targeted items is Rubén Blades. Four other sentences in the passage have him as their grammatical subject, three present tense and one periphrastic future tense. Learners might well have predicted during the test that if information referred to Blades, then they should simply indicate past, whether or not they had detected the lexical or the grammatical cue during processing. The mean scores on this test range from 42% to 68% to 80% for first, third and fifth semester learners, respectively, suggesting that such a discourse-based prediction strategy was not being employed. Had such a strategy been used, the scores should have been higher across all three levels of language learners. That kind of strategy would not be sensitive to linguistically-based developmental differences so that even first semester learners should have performed at the level of fifth semester learners had the strategy been employed.

Scoring procedures. In scoring the recalls, the number of instances where the learners correctly recalled a past temporal reference was tallied. Since the narrative contained seven verbs in the preterite tense, a total score of seven was possible. In the tense identification task, four items served as distractors, leaving seven test items. The recalls were scored independently by two researchers. Interrater reliability was .96.

Analysis. The scores obtained from the two dependent variables were submitted to a 3 x 2 Multivariate Analysis of Variance (MANOVA). The independent variables were language experience and input condition. The dependent measures were the scores on the reconstruction of propositional content and the tense identification test. We selected the MANOVA procedure since the two dependent measures were scored for the same seven targeted items.

Results

The MANOVA generated two separate Analyses of Variances (ANOVA), one for each dependent variable. The two ANOVAs will be discussed separately.

Reconstructing Propositional Content

The means and other descriptive statistics for the number of targeted past tense references appearing in the recalls are presented in Table 1. The means increase as semester level advances with the greatest difference occurring between first and third semester. A difference is also noted between the means grouped by input condition. Three

1011

times the number of past tense references were reconstructed when learners had lexical cues to tense as when they did not. To determine if these differences were significant, the means were submitted to an ANOVA, the results of which are presented in Table 2.

The ANOVA conducted on the recall task yielded significant main effects for both language experience as well as for input condition. There was no significant interaction between the independent variables.

TABLE 1
Means for Language Experience and Input Condition on the Reconstruction of Past Tense References

Semester Level	Mean	Std. Dev.	Std. Error
First	.289	.802	.130
Third	1.286	1.250	.211
Fifth	1.690	1.466	.272
Input Condition			
With Adverbs	1.491	1.502	.206
Without Adverbs	.531	.819	.117

TABLE 2
ANOVA of the Free Written Recall Task

Source	df	Sum of Squares	Mean Square	F-value	p-value
Language Experience	2	35.182	17.591	15.411	.0001
Input Condition	1	23.161	23.161	20.291	.0001
Language Experience X Narrative Condition	2	4.291	2.145	1.880	.1582

A post-hoc Scheffe's test revealed that both third and fifth semester learners reconstructed significantly more past temporal references than first semester learners ($p = .0007$ and $p = .0001$, respectively). However, there was no significant difference in the number of correct reconstructions of past temporal references between third and fifth semester learners ($p = .3262$). Another post-hoc Scheffe's test was performed on the means for input condition. This test demonstrated that those who listened to the narrative with lexical cues reconstructed significantly more past temporal references than those who only had grammatical cues in the input ($p = .001$).

Mean scores for the interaction between language experience and input condition are presented in Table 3. By looking at the numbers in the table of means, it is easy to see why no significant interaction was obtained for language experience and input condition: for all semester levels, learners who listened to a narrative containing both time adverbials and verb morphology did better than subjects who listened to a passage containing only verb morphology. However, the means obtained for third and fifth semester groups reveal a difference in scores much greater than the means for the first semester group. Given the extremely low mean scores of the first semester group (.550 for Condition 1 and .000 for Condition 2), we decided to probe the interaction even though it did not reach statistical significance. A means comparison contrast revealed that third and fifth semester learners recalled significantly more of the past instances under Condition 1 (adverbs and verb inflections) than under Condition 2 (verb inflections only). However, there was no effect for input condition for first semester learners. That is, they did not reconstruct significantly more past tense events under Condition 1 compared to Condition 2.

TABLE 3
Means for the Interaction Between Language Experience and Input Condition on the Free Written Recall Task

Semester Level	Input Condition	
	With Adverbs	Without Adverbs
First	.550	.000
Third	2.059	.556
Fifth	2.062	1.231

Tense Identification

The means and other descriptive statistics for the number of correct tense identifications of past temporal reference are presented in Table 4. The means increase as semester level increases with the greatest difference occurring between first and third semesters. A small difference is also noted between the means grouped by input condition. To determine if these differences were significant, means were submitted to an ANOVA, the results of which are presented in Table 5.

TABLE 4
Means and Other Descriptive Statistics for Language Experience and Input Condition on the Tense Identification Task

Semester Level	Mean	Std. Dev.	Std. Error
First	3.368	1.667	.270
Third	4.743	1.268	.214
Fifth	5.276	1.009	.204
Input Condition			
With Adverbs	4.623	1.643	.226
Without Adverbs	4.122	1.523	.218

TABLE 5
ANOVA of the Tense Identification Task

Source	df	Sum of Squares	Mean Square	F-value	p-value
Language Experience	2	65.623	32.812	17.174	.0001
Input Condition	1	6.515	6.515	3.410	.0679
Language Experience X Narrative Condition	2	1.671	.836	.437	.6470

The ANOVA conducted on the tense identification task yielded a significant main effect for language experience. However, there was no significant main effect for input condition and there was no significant interaction between language experience and input condition.

A post-hoc Scheffe's test revealed a pattern similar to that on the reconstruction data. Both third and fifth semester learners identified significantly more past temporal references than first semester learners ($p = .0003$ and $p = .0001$, respectively). However, there was no significant difference in the number of correct identifications of past temporal references between third and fifth semester learners ($p = .3120$). In other words, learners with greater language experience (in this case third and fifth semester learners) identified significantly more of the past instances than lower level learners (first semester learners). Although input condition did not reach a level of statistical significance, a trend toward performance enhancement emerged when lexical cues were present ($p = .0679$).

Mean scores for the nonsignificant interaction between language experience and input condition are presented in Table 6. The means show a consistent pattern: learners at the three levels of language experience consistently identified more of the past references when they listened to a narrative that contained both adverbs and verb inflections. Unlike the probing conducted on the recall scores, however, a means comparison contrast did not reveal any significant differences on the tense identification task.

TABLE 6
Means for the Interaction Between Language Experience and Input Condition on the Tense Identification Task

Semester Level	Input Condition	
	With Adverbs	Without Adverbs
First	3.650	3.056
Third	4.824	4.667
Fifth	5.625	4.846

1015

Discussion and Conclusion

This study was undertaken to examine the effects of lexical and grammatical cues on processing past temporal reference. It adds to the growing body of research on attention, input processing, and the construction of L2 linguistic systems through processing input by identifying formal features of the language that manage to get registered in memory. Once registered in memory, which features become available for other cognitive processing? Learners in this study listened to one of two versions of the input passage. One group heard a narrative that contained seven targeted past temporal references encoded morphologically and with lexical cues to temporal reference. Another group heard the same narrative with the same targeted items. The lexical cues were, however, deleted from the sentences in which the targeted references occurred. After listening, subjects first had to reconstruct the propositional content in order to determine what they actually processed and stored as the tense of particular events in the narrative and then had to indicate the tense of particular events mentioned in the narrative. Subjects from three levels of language experience were used in order to determine whether a developmental pattern would emerge.

Previous research on discourse level processing indicates that orienting attention toward formal features of language yields a loss of propositional content whereas orienting attention toward a semantic element of the discourse yields no such loss (Bransdorfer, 1991; Hulstijn, 1989; VanPatten, 1990). Our study did not orient learners to one feature of the input or another but focused them to listen for meaning. Rather, we manipulated the characteristics of the input as has been done in research on sentence level processing (e.g., Musumeci, 1989). In the absence of differing orienting instructions, learners reconstructed more propositional content when the input contained lexical cues than when the input contained only grammatical cues to temporal reference. Learners aligned attention on the lexical cues and better utilized them to reconstruct propositional content. The grammatical cues obviously received some attention but were not as useful in reconstructing propositional content.

Scores on the tense identification task were much higher than those on the reconstruction of content. Previous research on comprehension assessment supports this finding in that prompted response tasks yield higher scores than more open-ended tasks (Berne, 1992; Lee, 1987; Shohamy, 1984; Wolf, 1993). Whereas the first semester learners reconstructed almost no past temporal references, they did identify 42% of the past temporal references when prompted. If we had used only

1016

one measurement, the reconstructions, we would have had to conclude that first semester learners found the input incomprehensible. Based solely on the performance of first semester learners on the reconstruction tasks, we concluded that learners with less than sixty contact hours with the target language can not detect the temporal cues in the input, be they lexical or grammatical, in such a way that the cues become available for other cognitive processing. The combination of measurements allows us, however, to say that first semester learners are indeed noticing the forms in the input. They may not yet be able to utilize what they notice to express meaning, but the evidence suggests that they are detecting some features in the input. It may also be the case as Clarke (1979) suggests, that the first semester learners had not attained a sufficient level of L2 proficiency to perform well on the comprehension assessment; limited proficiency short-circuits comprehension. Similarly, VanPatten's principles maintain that as learners gain in proficiency they are more able to attend to both propositional content and grammatical form. Beginners, such as these first semester learners, would be able to do one or the other.

The results of the tense identification task also showed that the effect of input condition was leveled out across all three levels of learners; independent of proficiency level, learners did not identify significantly more past temporal references when they had the lexical cues in the input than when they did not. The structure and prompted nature of the identification task may account for this finding. Lexical cues were not a part of the prompts for either group of learners; they all, despite the input condition to which they were assigned, took the same tense identification task consisting of the infinitive form of a verb plus other elements of the verb phrase,

Another account of this finding is that the tense identification task indicates that learners, especially the more advanced ones, do attend to and detect grammatical cues when they are processing for meaning (a notion consistent with VanPatten's principles). These grammatical cues may not, however, be that useful to the learners in reconstructing propositional content, hence the advantage lexical cues give learners in reconstructing content. When given a piece of information from the text as a prompt, more advanced learners can identify temporal reference. But when they have to access the meaning stored in memory, grammatical cues are less helpful than lexical ones.

The study revealed a consistent pattern of developmental differences. Across both dependent variables the first semester learners did not perform to the same level as the third and fifth semester groups; these latter two exhibited no quantitative differences between them.

1017

Much is being written on consciousness (Ellis, 1994; Schmidt, 1993) and consciousness raising (Fotos, 1993; Harley, 1994; Sharwood Smith, 1993). The results of the present study also contribute to the discussion of directing learners' attention toward formal features of the input in instructed settings. We can certainly question whether learners who assign meaning to the lexical cue to past temporal reference ever then process and make use of the grammatical cue. Previous empirical research suggests strongly that they do not (Peters, 1985; Sharwood Smith, 1985, 1993; Tomlin & Villa, 1994; VanPatten, 1990). Since all learners in this experiment were focused on deriving meaning from the input, their processing strategies were self-directed toward features of the input that would aid them in creating meaning. As Tomlin and Villa's research on attention indicates, once the lexical cue is detected it would interfere with the detection of other information, such as a grammatical cue, since the lexical cue consumes attentional resources.

From a pedagogical perspective, we can ask then, how learners' attention might be directed toward formal features of the input so that they process them. That is, how can learners be directed both to make meaning and to make form-meaning connections? A type of grammatical instruction called processing instruction investigates the connection between input processing, comprehending input, and building linguistic systems. The research carried out to date, summarized in VanPatten (1996), consistently reports the benefits of grammatical instruction aimed at having learners attend to formal features of the input provided they attach meaning to the form (Cadierno, 1995; Cheng, 1995; VanPatten & Cadierno, 1993; VanPatten & Sanz, 1995). In these studies, learners not only gain in their ability to comprehend grammatical form during input processing, but they also gain in their ability to use the form in output. Both theory and pedagogy have something to gain by a continued investigation of how learners attend to input data.

References

Barasch, R. M., & James, C. V. (Eds.). (1994). *Beyond the Monitor Model: Comments on current theory and practice in second language acquisition.* Boston: Heinle & Heinle.

Bardovi-Harlig, K. (1992). The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics, 13,* 253-278.

1018

Berne, J. (1992). The effects of text type, assessment task, and target language experience on foreign language learners' performance on listening comprehension tests. Unpublished doctoral dissertation, The University of Illinois at Urbana-Champaign.

Bransdorfer, R. (1991). Communicative value and linguistic knowledge in second language oral input processing. Unpublished doctoral dissertation, The University of Illinois at Urbana-Champaign.

Ellis, R. (1994). Consciousness in second language learning: Psychological perspectives on the role of conscious processes in vocabulary acquisition. *AILA Review, 11*, 37-56.

Fotos, S. (1993). Consciousness raising and noticing through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics, 14*, 385-407,

Gass, S. M., & Madden, C. G. (Eds.). (1985). *Input in second language acquisition*. Cambridge, MA: Newbury House.

Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9*(2), 198-217.

Harley, B. (1994). Appealing to consciousness in the L2 classroom. *AILA Review, 11*, 57-68.

Hulstijn, J. (1989). Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H. Dechert & M. Raupauch (Eds.), *Interlingual processes*, 49-73. Tübingen: Gunter Narr Verlag.

James, W. (1890). *The principles of psychology*. New York: Holt.

Klein, W. (1986). *Second language acquisition*. Cambridge, UK: Cambridge University Press.

Klein, W. (1994). Learning how to express temporality in a second language. *Societa di Linguistica Italiana, 34*, 227-248.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Larsen-Freeman, D., & Long, M. H. *An introduction to second language acquisition research*. London: Longman.

Lee, J. F. (1987). Comprehending the Spanish subjunctive: An information processing perspective. *Modern Language Journal, 71*, 50-57.

Lee, J.F., & VanPatten, B. (1995). *Making communicative language teaching happen*. New York: McGraw-Hill.

Leow, R. (1995). Modality and intake in second language acquisition. *Studies in Second Language Acquisition, 17*, 479-489.

Long, M. H. (1991). Focus on form: A design feature in language teaching. In K. de Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp.39-53). Amsterdam: Benjamins.

Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16*, 283-323.

Mangubhai, F. (1991). The processing behaviors of adult second language learners and their relationship to second language proficiency. *Applied Linguistics, 12*, 268-297.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second-language learning: An information processing perspective. *Language Learning, 33*, 135-158.

Musumeci, D. (1989). The ability of second language learners to assign tense at the sentence level. Unpublished doctoral dissertation, The University of Illinois at Urbana-Champaign.

Peters, A. M. (1985). Language segmentation: Operating principles for the perception and analysis of language. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition, 2* (pp. 1029-1067). Hillsdale, NJ: Lawrence Erlbaum.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Schmidt, R. (1993). Awareness and second language acquisition. *ARAL, 13*, 206-226.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Sharwood Smith, M. (1985). From input to intake: On argumentation in second language acquisition. In S. M. Gass, & C. G. Madden (Eds.). *Input in second language acquisition* (pp. 394-403). Cambridge: Newbury House.

Sharwood Smith, M. (1986). Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics, 7*, 239-274.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*, 165-179.

Slobin, D. (1985). *The crosslinguistic study of language acquisition: Theoretical issues, 2*. Hillsdale, NJ: Lawrence Erlbaum.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G.Madden (Eds.), *Input in second language acquisition* (pp. 235-256). Cambridge, MA: Newbury House.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183-203.

VanPatten, B. (1985). Communicative value and information processing in second language acquisition. In P. Larson, E. Judd, & D. Messerschmidt (Eds.), *On TESOL '84: A brave new world for TESOL* (pp. 89-100). Washington, DC: TESOL.

VanPatten, B. (1989). Can learners attend to content and form while listening to the L2? *Hispania, 72*, 409-417.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquistion, 12, 287-302.*

VanPatten, B. (1992). Second language acquisition research and foreign language teaching, Part 2. *ADFL Bulletin, 23*(3), 23-27.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

Wolf, D. (1993). A comparison of assessment tasks to measure foreign language reading comprehension. *Modern Language Journal, 77*, 473-488.

Appendices

Appendix A

Narrative with Lexical and Grammatical Cues

¡Hola amigos! Soy José Pérez del canal 8 de Radio Televisión Panameña. Hoy vamos a hablar de Rubén Blades, famoso cantante de salsa panameño.

La semana pasada, yo conversé con él en su casa y él me habló de sus preocupaciones sociales y políticas. Por ejemplo, es importante mencionar su contribución al video y disco *Sun City* contra el Apartheid. También, *una vez en el pasado*, expresó la idea de traer la democracia a Panamá, Con respecto a su carrera musical, Rubén Blades participó en varias bandas de música latina *durante los años 60. En 1976* trabajó con Willie Colón en una serie de discos. El más importante se llama *Salsa*. La canción más popular en este disco se llama "Pedro Navaja." Esta canción cuenta la historia de un crimen en un barrio de una ciudad.

Pero, querido público, Rubén Blades no es solamente músico. *Hace cuatro años* decidió volver a la universidad para hacer un Master en leyes y ahora va a trabajar en el cine también. *Ayer* me llamó para

1021

contarme esto. Tiene una oferta para trabajar en Hollywood. Como pueden ver, Rubén Blades es una persona con múltiples intereses. Obviamente, parece que su carrera va muy bien.

Narrative Without Adverbs

¡Hola amigos! Soy José Pérez del canal 8 de Radio Televisión Panameña. Hoy vamos a hablar de Rubén Blades, famoso cantante de salsa panameño.

Yo conversé con él en su casa y él me habló de sus preocupaciones sociales y políticas. Por ejemplo, es importante mencionar su contribución al video y disco *Sun City* contra el Apartheid. También expresó la idea de traer la democracia a Panamá. Con respecto a su carrera musical, Rubén Blades participó en varias bandas de música latina. Trabajó con Willie Colón en una serie de discos. El más importante se llama *Salsa*. La canción más popular en este disco se llama "Pedro Navaja". Esta canción cuenta la historia de un crimen en un barrio de una ciudad.

Pero, querido público, Rubén Blades no es solamente músico. Decidió volver a la universidad para hacer un Master en leyes y ahora va a trabajar en el cine también. Me llamó para contarme esto. Tiene una oferta para trabajar en Hollywood. Como pueden ver, Rubén Blades es una persona con múltiples intereses. Obviamente, parece que su carrera va muy bien.

Appendix B

Tense Identification Test

Instructions:

The following phrases refer to events in the passage you just heard. Indicate whether the action occurs (present), occurred (past) or will occur (future):

*1. TENER una oferta de Hollywood
   a. present
   b. past
   c. future

2. PARTICIPAR en bandas de música latina
   a. present
   b. past
   c. future

1022

3. CONVERSAR con José en su casa
   a. present
   b. past
   c. future

4. VOLVER a la universidad
   a. present
   b. past
   c. future

5. LLAMAR por teléfono
   a. present
   b. past
   c. future

6. HABLAR de sus preocupaciones políticas y sociales
   a. present
   b. past
   c. future

*7. HABLAR de Rubén Blades en la radio
   a. present
   b. past
   c. future

8. EXPRESAR la idea de traer la democracia a Panamá
   a. present
   b. past
   c. future

9. TRABAJAR con Willie Colón en una serie de discos
   a. present
   b. past
   c. future

*10. TRABAJAR en el cine
   a. present
   b. past
   c. future

*11. IR su carrera muy bien
   a. present
   b. past
   c. future

* Indicates a distractor

Authors

TERESA CADIERNO, Associate Professor, Spanish, Oden University, Campusvej 55, DK 5230 Odense M., Denmark. *Specializations:* acquisition of Spanish, effects of instruction.

WILLIAM R. GLASS, Sponsoring Editor for Foreign Languages, The McGraw Hill Companies, 55 Francisco St., San Francisco, CA 94133. *Specializations:* second-language reading research, second language teaching.

BILL VANPATTEN, Professor, Spanish, Italian, and Portuguese, University of Illinois at Urbana-Champaign, 4080 FLB MC-176, 707 S. Mathews Ave., Urbana, IL 61801. *Specializations:* input and input processing in second language acquisition, the effects of instruction.

JAMES F. LEE, Associate Professor, Spanish and Portuguese, Indiana University of Illinois, Ballantine 802, Bloomington, IN 47405. *Specializations:* second-language reading research, classroom language acquisition.

1024

Applied Language Leaning
1997, Vol. 8, No. 1, pp. 25-42

A Link Between Reading Proficiency and Native-Like Use of Pausing in Speaking

Ruth Johnson and Rita Moore
Southern Illinois University

The research reported in this article investigated whether there is a correlation between a nonnative English speaker's reading proficiency in English and the use of native-English-like pausing in reading aloud. Seventy-six English-as-a-Second-Language (ESL) students in high intermediate and advanced classes were given two tests, one a read-aloud passage and the other a reading test. The reading test consisted of the reading section of a Test of English as a Foreign Language (TOEFL) examination, whereas the read-aloud passage was a paragraph that included numerous complex sentences, words in series, and dependent clauses. Results indicate a statistically significant although moderate correlation between students' reading test scores and their native-like use of pausing in reading aloud. Further research is recommended to determine whether or not there exists a cause-and-effect relationship, with a view to making recommendations for what to teach in reading and speaking classes, vis-à-vis thought groups and breath groups.

Although speech contains many characteristics, when native speakers (NSs) react to nonnative pronunciation, they most often do so in terms of overall comprehensibility and acceptability (Ludwig, 1982; Johansson, 1978). Although NSs do not attempt to count or measure various types of errors, such as errors in the use of segmentals, deviance in prosody (suprasegmentals), and nonnative-like syllable structure errors, they are aware of their negative impact on the communication process. The question of the relative importance of these different components of nonnative pronunciation for both comprehensibility (Anderson-Hsieh, Koehler, & Johnson, 1992) and acceptability (Delamere, 1986; Johnson, 1993; Wang, 1994) has been explored. Many in applied linguistics have emphasized the importance of understanding the roles

that elements of speech play in comprehensibility and acceptability. The prominence of the suprasegmentals, among them, pausing, in the intelligibility of speech (Hatch, 1983; Pennington & Richards, 1986; Wong, 1987; Acton, Gilbert, & Wong, 1991) has been stressed. The suprasegmentals are critical because, in a stress-timed language such as English, they direct the listener's attention to information the speaker regards as important (Gutknecht, 1991).

For an English NS listening to speech produced by a nonnative speaker (NNS), there is what can be called a "threshold of comprehensibility." At the level below this threshold, speech is incomprehensible or extremely difficult to comprehend (Anderson-Hsieh, et al., 1992; Johnson, 1990). The causes of difficulty in the speech signal are many. The speech signal may contain many segmental errors on both the phonetic and the phonemic levels (Flege, 1981; Beebe, 1987). Further, it may deviate from NS pausing, stress, intonation, and rhythm, so that, for example, the speech may have a pronounced sing-song intonation or syllable-timed rhythm (Dauer, 1983; Bolinger, 1986; Tarone, 1987a). Also, the speech signal may contain numerous syllable structure errors, usually in consonant deletion and vowel insertion (Broselow, 1987; Sato, 1987; Tarone, 1987b) or it may contain many grammatical errors (Delamere, 1986).

However, when the threshold of comprehensibility has been reached, research has shown that improvement in speech (accent reduction) is greater when the suprasegmentals are improved than when errors in segmentals are reduced (Anderson-Hsieh, et al., 1992; Hieke, 1988). The suprasegmentals, however, are difficult to teach and difficult for NNSs to master, and they traditionally have been given little attention in speaking or pronunciation classes. NNSs, numbering about 25, who have been students in an "Accent Reduction" class taught by these researchers (unrelated to the subjects and classes involved in this study), also have reported, in exit interviews for that class, that lessons in the use of the suprasegmentals for accent reduction lack face validity. In other words, they value pronunciation lessons that focus on segmental errors; for example, a lesson that would drill them in the difference in the "th" sound and the "d" or "t" sound is seen as useful. To focus on the "music" of the language, its prosody, the suprasegmentals, is viewed as a waste of time.

Some of the elements involved in teaching and learning the use of breath groups in English include the use of appropriate pausing. Chafe (1988) reports that "spoken language exhibits important prosodic units...(called) 'intonation units'" (p. 397). This looks remarkably like some of the elements that are taught when one teaches and learns

reading in English, namely, the ability to organize one's reading into thought groups, using, for example, punctuation (written "pausing"). Long (cited in Chafe, 1988) describes the "auditory imagery," based upon mental pausing, that would be apparent as prosody if written discourse were to be read aloud. Chafe (1988) maintains that there is a link between prosody and punctuation, although "people who read aloud nearly always produce intonation units whose length lies within the normal range for ordinary spoken language" (p. 424), that is, shorter than the processing length for reading.

Another issue involved in second language learning is acquiring the knowledge of immediate constituent grammar (Radford, 1981). This concept holds that, because a NS possesses intuitive knowledge of English syntax, the relationships between the parts of a sentence are known. Thus, the sentence, "The man ate the sandwich," can be divided into two groups: "The man..." and "...ate the sandwich" and cannot be divided into these two groups: "The man ate..." and "...the sandwich" because a rule in immediate constituent grammar is that the direct object is more closely tied to the verb than is the subject to the verb. Hence, "The man..." is considered to be a structural constituent because these two words go together in a group (Radford, 1981). The ability to recognize such constituent groups could play a part in the ability of the second-language learner (SLL) to read and speak in thought/breath groups because constituent groups can comprise thought/breath groups and constituent groups of words are not to be interrupted within the group.

Some research has addressed the issue of the reader's ability to read in thought groups. Alderson (cited in Alderson & Urquhart, 1984) hypothesizes that people learning a second language (L2) may have difficulty reading in the L2 because they employ "incorrect strategies for reading that foreign language" (p. 4); in other words, in the case of ESL, they may not know how to chunk properly. Alderson (cited in Alderson & Urquhart, 1984) says that there is evidence that "poor foreign language reading is due to incorrect strategies for reading that foreign language, strategies which differ from the strategies for reading the native language" (p. 4).

Just and Carpenter (1980) found that the reader's knowledge of the "topic, syntactic constraints, and semantic associates...plays a role in activating and selecting the appropriate concepts" (p. 352). Reading ability may depend on "the automaticity of basic reading processes such as encoding and lexical access" (p. 351), which are functions of one's proficiency in a language.

1027

Related research has investigated the notion of chunking which is the surface characteristics of text that are used to aid comprehension and that are dependent upon sentence and phrase boundaries. Levelt (1993) has constructed a language processing model in which a speaker passes through a microplanning step in which "chunks" of language are assigned propositional shape and informational perspective. The speaker's ability to chunk language so that it is more readily accessible to the listener may have a parallel in one's ability to read language in chunks and process it more readily.

Kintsch and Van Dijk (1978) state that "a good speaker attempts to place his or her sentence boundaries in such a way that the listener can use them effectively for chunking purposes" (p. 368). Chunks are used to package information in a suitable size for short-term memory to realize a cyclical comprehension process; those with "low verbal abilities" cannot access information in short-term memory as readily as can those with better verbal abilities (Kintsch & Van Dijk, 1978). Higher reading abilities could also be dependent upon the ability to process chunks effectively; if SLLs do not chunk well, they may not be able to read as well.

Miccinati (1985) indicates that children learning to read in their first language (L1) rely heavily on prosodic cues. Children need to be trained to recognize prosodic features in language, becoming "aware of the predictable language patterns of syntax" as they become "sensitive to...prosodic signals" (Miccinati, 1985). Thus, learning to read by employing knowledge of the suprasegmentals of a language can help the reader comprehend the meaning of the text.

Meaning of text is also related to a reader's ability to use parsing skills: the ability to group sentences into meaningful phrases (thought groups), which aids language comprehension. Flippo (1984) found that children who have reading problems often lack such skills. When they are made sensitive to prosodic signals, then they are able to *predict* language patterns of syntax and, consequently, read better.

Other studies have looked at the relationships between second language oral proficiency and reading comprehension. Peregoy (1989) reported a correlation between L2 oral performance and reading performance on standardized texts, that is "reading for meaning." This "meaning" construct would include the ability to read in thought groups; that is, the ability to group sentences into meaningful phrases reported by Miccinati (1985). Beyond the correlational relationship, Peregoy (1989) found that oral proficiency limits L2 reading ability.

Preliminary results on qualitative research currently being conducted by the authors, using as subjects international teaching assistants who

1028

want to improve their speaking, suggests that a close relation exists between L2 speaking and L2 reading, coinciding in development over time. Also, one of the teaching assistants, Ballard, reported to these authors that an ESL student remarked to her teacher that when *she herself* noticed improvement in her use of the suprasegmentals that she realized her *reading ability* had improved (personal communication, 1994); she said that the language "in her head" was organized into phrases that matched, for example, the intonation contours that she had learned in pronunciation classes.

Reading should be done with correct phrasing and intonation (Dowhower, 1991). *Prosodic reading* is the ability to read in expressive rhythmic and melodic patterns. Fluent readers are able to organize text into meaningful units, including appropriate pausing, long phrasing and phrase-final lengthening, elimination of pausal intrusions, correct stress, and correct terminal intonation contours (Browne & Huckin, 1987). Fluent speakers are able to do the same (Anderson-Hsieh, et al., 1992). Oral reading, which combines the abilities to read with the abilities to speak, can be used to determine the ability of an SLL to read and speak using thought groups (Miccinati, 1985). In addition, Daneman and Carpenter (1983) found that "the similarities between silent and oral reading...are more striking than the differences" (p. 579). In both, immediate integration processes include the detection and resolution of inconsistency. Contradictions (or seeming contradictions) cause readers, both silent and oral, to pause; these pauses occur at boundaries (Daneman & Carpenter, 1983).

In oral skills, Taylor (1981) states that uneven, jerky rhythm results "from faulty division into sense and breath groups" (p. 237); that is, the point where the NNS of English takes a breath that does not coincide with what makes sense semantically to the NS listener. The NNS is not respecting "utterance units," those units of speech that occur between identifiable (to an NS) pauses or breaks in tempo (Maynard, 1986).

In speaking, if boundaries are marked inappropriately by the NNS, the result is difficulty on the part of the native speaker-listener to follow the thread of speech (Wennerstrom, 1994). The ability to pause properly is actually a mark of authentic-sounding speech (Scanlan, 1987). Gutknecht (1978) maintains that "the importance for communication lies primarily in the *attitudinal* function of intonation. That is, in conversation with a foreigner, an Englishman is prepared for grammatical and lexical mistakes, but in general automatically interprets mistakes in intonation as being a reflection of the wrong attitude (on the part of the NNS)" (emphasis added) (p. 259-260). The same can occur with errors in pausing.

The purpose of the present study is to investigate whether there is a correlation between NNSs reading proficiency in English, as measured by the score they obtain on the TOEFL exam, section III (reading comprehension section) and their use of correct pausing in English while reading aloud. Reasons for choosing a read-aloud passage rather than spontaneous speech are outlined below in the "Method" section.

Method

A sample of ESL students from an intensive English program were given two tasks, one a read-aloud passage and the other a standardized reading test (reading section of TOEFL). The subjects for this study were asked to read aloud a short passage which was recorded and analyzed for nonnative-like pausing.

The Read-Aloud Task

Such task was deemed appropriate for use in this study for several reasons. First, a read-aloud task allowed for control over the content of the speech and the predictability of the thought/breath groups. In the semester preceding the experiment, the passage was piloted on a Level 3 group of Southern Illinois University students in the ESL Center, who were comparable to the experimental group; the passage was determined to be easily readable. In other words, none of the vocabulary items was unknown, and sentence length and complexity were within the English-proficiency level of the subjects. Second, the passage was chosen for its predictability for pausing. It was written in sentences that contained short breath groups, ranging in number of words from three to twenty, with 17 of the 21 breath groups in the passage containing ten or fewer words (the other four contained 12, 13, 18, and 20 words).

To predetermine that the breath groups were "predictable," the passage was given to six native readers to read aloud. These native-reader controls were undergraduate students at Southern Illinois University at Carbondale, the same university where the NNSs were attending an intensive English program. None of the controls paused at a place other than that predetermined by the researchers to be a juncture between breath groups. Also, none of them broke the passage into 21 breath groups by breathing; instead, they relied on intonation contours to define breath groups. For the purposes of this study, however, the researchers accepted a pause between any two breath groups even if the NNS actually breathed 21 times while reading the passage *so long as the breath was taken at the breath-group junctures*. A pause (breath) taken

1030

within the predetermined breath groups was defined as a "nonnative-like pause." Also, for the purposes of this study, the researchers accepted one compensation strategy which was evident in two of the NSs: They would pause at an inappropriate place but then, as they began speaking again, *they would back up to the beginning of the thought/breath group and repeat the phrase with native-like pausing.* This strategy we accepted as native-like when used by our subjects.

Subjects

A group of 76 Levels 3-4 (intermediate-advanced) students in the ESL Center at Southern Illinois University at Carbondale were asked to participate in this study as a part of their classroom instruction. There were 42 students in Level 3 and 34 students in Level 4; 40 were female and 36 were male; 20 different countries were represented and 10 languages: Thai (19), Chinese (13), Korean (12), Japanese (10), Spanish (10), Arabic (6), Indonesian (2), Greek (2), Turkish (1), and Portuguese (1).

Tasks: Type and Administration

The subjects were given the two tasks as a part of their regular classroom instruction. Both tasks were administered in the language media center, one after the other; the read-aloud test was audiotaped. The students were told that they were part of a research project investigating reading skills and speaking skills (while reading aloud) in English.

The first task was the reading portion of the TOEFL test, which was administered according to Educational Testing Service (ETS, Princeton, New Jersey) rules for that portion of the test. The second task was an individual read-aloud test in which the subjects were asked to read aloud a passage that they had not seen before; this reading was tape-recorded.

Variables

With reference to reading aloud, the specific variable that was investigated was the ability to speak using correct breath groups. The scores obtained on this test, which were measured as the number of times an *incorrect* (nonnative-like) pause was made, were correlated with the students' reading proficiency test scores. A correlation was deemed appropriate for this study because prior to this study no work had been done on the relationship between reading and organizing

1031

oneself in speaking using thought/breath groups. If a correlation would be established, then a study using an analysis of variance or a multiple regression test would be of interest to determine predictability of one variable on the other.

Scoring

The tape-recorded read-aloud passage was played for two NS scorers who counted the number of incorrect pauses made by each student. Inter-scorer agreement on the scores for the read-aloud passage was calculated by dividing the number of agreements by the number of agreements plus disagreements. An agreement was any inter-word or inter-sentence position at which both scorers scored a nonnative-like pause. Disagreements were those instances where one scorer scored a nonnative-like pause and the other did not. The average level of agreement was 91%.

Results and Discussion

Overall Relationship

The first analysis was conducted to reveal the relationship between reading proficiency in English and the students' ability to read aloud using native-like pausing. For purposes of this analysis, reading proficiency in English was operationally defined by the students' scores on the reading section of the TOEFL test. The students' read-aloud ability was operationally defined by the number of nonnative pauses they made while reading aloud a passage of approximately one minute in duration. Pearson product-moment correlations were computed to determine the strength of the relationship between reading measures and performance on reading aloud.

The means and standard deviations for the students' performance as an entire group on the reading test and the read-aloud task are presented in Table 1. The read-aloud mean scores given in Table 1 and Table 2 are the mean number of *incorrect* pauses (nonnative-like pauses) for the entire sample. Scores for the reading test are reported as standard scores for the reading section of the TOEFL exam.

Also given are the means and standard deviations for the students' performance by L1 writing system (Table 2). These data are provided because, as Bernhardt (1992) has suggested, grouping ESL students together when measuring reading or writing proficiency, without regard to a comparison of their writing systems with English, may be invalid.

What is of interest in these numbers is the fact that native writers of Arabic, Chinese, and Japanese did differ from native writers of Spanish, the latters' scores being meaningfully different from the scores of the former groups, as would be expected since the Spanish writing system is the closest of all the groups. What is not so easily explainable is the performance of the Korean and Thai students who did almost as well as the Spanish students, all of whom performed above the mean in both the reading test and the read aloud test. However, what may explain this difference is exactly what the research question of this report is: that reading ability and reading aloud ability are closely linked and are more a function of one's interlanguage and L2 abilities than a function of L1 interference.

TABLE 1
Descriptive Statistics for Reading Proficiency Test and Read-Aloud Task for Entire Sample (N=76)

Test	Min	Max	Mean	Median	Mode	Std Dev
Reading	28	58	46.05	48.00	48.00	6.85
Reading Aloud	1	29	11.03*	10.00	5.00	7.73

*This number, as a mean, is reported as a positive number. However, its correlation with the mean of "Reading" will be negative because the "Reading" score measures proficiency in reading while the "Reading Aloud" score reflects errors made in pausing while reading aloud.

The primary question was the relationship of the students' reading ability in English, as measured by the standardized reading test, and their ability to use correct pauses (breath groups) when reading aloud. A correlation analysis was conducted; results indicated a moderate and statistically significant correlation of -0.3497 ($p<.002$). The negative correlation was expected because the scores were compared to a score on the standardized reading test (a positive integer), on the one hand, and a score indicating the number of *incorrect* pauses (a negative score) on the other hand. The moderate, significant correlation is important: NNSs' ability to read and their ability to process information into breath groups, as measured by pausing, were related somewhat.

TABLE 2
Descriptive Statistics for Reading Proficiency Test and Read-Aloud Task for Subgroups by Native Writing System

Test	Min	Max	Mean	Median	Mode	Std Dev
Arabic (N=8)						
Reading	28	48	40.12	41.50	43.00	5.90
Reading Aloud	5	29	18.75*	14.00	14.00	14.35
Chinese (N=13)						
Reading	29	51	39.62	40.00	40.00	6.55
Reading Aloud	5	24	14.46*	15.00	15.00	5.87
Japanese (N=10)						
Reading	35	52	42.60	42.50	40.00 45.00 48.00	5.78
Reading Aloud	5	19	13.50*	13.50	9.00 17.00	4.99
Korean (N=12)						
Reading	36	58	48.25	48.50	48.00	6.74
Reading Aloud	4	20	10.42*	8.50	7.00	5.33
Spanish (N=11)						
Reading	47	58	53.45	53.00	53.00 57.00	3.16
Reading Aloud	1	13	7.64*	7.00	5.00	3.67
Thai (N=19)						
Reading	43	55	49.21	48.00	48.00	2.90
Reading Aloud	2	23	8.26*	6.00	10.00	8.26

*These numbers, as means, are reported as positive numbers. However, their correlation with the means of "Reading" will be negative because the "Reading" score measures proficiency in reading while the "Reading Aloud" score reflects errors made in pausing while reading aloud.

Additionally, the standardized regression analysis (R2) amounted to 0.12, indicating that 12% of the variation in the total of nonnative pauses (the dependent variable) can be associated with changes in the reading test score (the independent variable). Although this percentage

may appear modest, it is meaningful because the analysis was done with errors in pausing on a holistic reading score. In other words, this figure would indicate that out of all the variables that contribute to or affect reading, errors in pausing account for 12% of that total.

Effect of Class Level on Relationship Between Reading Ability and Ability to Use Native-Like Pausing

A secondary analysis was conducted to reveal the effect that the students' class level within the intensive English program had on the relationship between reading proficiency in English and their ability to read aloud using native-like pausing. Pearson product-moment correlations were computed to determine the strength of the relationship between reading measures and performance on reading aloud when class level was controlled.

The question to be addressed was the effect of the students' attainment of English, as measured by their placement in an intensive English program (determined by a TOEFL score combined with teachers' recommendations), on the relationship of the students' reading ability in English, as measured by the standardized reading test, and their ability to use correct pauses (breath groups) when reading aloud. Correlation analyses were conducted using class level 3 (class size=42) and class level 4 (class size=34) as subgroups. A correlation analysis was conducted; results indicated a moderate and statistically significant correlation in each subgroup. For class level 3, the correlation was -0.3359 ($p<0.032$), and for class level 4, the correlation was -0.4672 ($p<0.005$). This finding is important: it shows that the effect of level of proficiency, as measured by class placement within an intensive English program, was significant vis-à-vis the relationship of the students' reading ability in English and their ability to read aloud in English using native-like pausing.

Relationship Between Reading Ability and Nonnative-Like Pausing Within Specific Environments

Another analysis was conducted to test the relationship between the students' reading ability and specific places at which they made nonnative-like pauses when reading aloud. For this analysis six "error environments" were selected; the criterion for choosing an error environment was determined by counting the potential number of times within the reading passage that that type of error could occur. For example, there were 16 instances within the passage in which a pause before a

prepositional phrase would be nonnative-like; there were 24 instances in which pausing between an adjective and the noun it modified would be nonnative-like. The error environments are presented in Table 3 and their relationship with reading ability is given in Table 4.

TABLE 3
Error Environments and Average Number of Nonnative Pauses

Environment	Total Envs	Total # Pauses	Average	# Ss Making Error
Between Subj. & Verb	7	16	2.3	14
At End of a Line	13	119	9.2	50
Between Verb & DO	10	63	6.3	35
Between Prep & Obj	15	60	4.0	38
Before Prep Phrase	16	136	8.5	55
Between Adj & Noun	24	124	5.2	61

Looking at specific environments, we expected that nonnative-like pausing would occur among lower-proficient readers as compared to higher-proficient readers to a significant degree in all of the selected environments:

1. at the end of a line
2. between an adjective and the noun it modified
3. between a preposition and its object
4. between a verb and its direct object
5. between the subject and the verb
6. before a prepositional phrase

1036

TABLE 4
Correlation of Reading Ability with Nonnative-Like Pausing when Reading Aloud in Selected Groupings of the Sample

Grouping	Correlation Value**
Entire Sample	-0.3512
By Class	
Class 3	-0.3259
Class 4	-0.4925
By Error Environment	
Between Subj. & Verb	*
Between Verb & DO	-0.3421
Between Prep & Obj	*
At the End of a Line	*
Before Prep Phrase	-0.2759
Between Adj & Noun	*

**significant at $p<0.05$
*not significant at $p<0.05$

When we analyzed the number of students who made an error in a given environment, we observed that 61 (80.3%) of the students paused between an adjective and the noun it modified, 55 (72.4%) paused before a prepositional phrase, and 50 (65.8%) paused at the end of a line. Because in each case these pauses involved at least 2 of every 3 subjects, one could predict that the level of students' reading proficiency would not result in a significant difference for these three environments.

For two of these environments this hypothesis was true: Pausing between an adjective and the noun it modified and pausing at the end of a line were done even as reading proficiency increased; there was no significance when a correlation of higher- and lower-proficient readers was done. On the other hand, there was a significant difference when reading ability was correlated with pausing before a prepositional phrase (the type in which the reader recognizes the beginning of a phrase which is integrally related to the phrase preceding it). On the whole, the correlation, however, was moderately weak: -0.2759. This type of error in pausing, that is, pausing at constituent boundaries, could very well become insignificant as reading proficiency increases, whereas pausing between an adjective and the noun it modifies and pausing at the end of a line, that is, within constituent pausing, persists despite the level of reading proficiency.

1037

Correlations could be expected in those environments in which roughly half of the students made a nonnative pause that is, pausing within an infinitive phrase, pausing between a preposition and its object, and pausing between a verb and its direct object. However, again, pausing in only one of these environments, namely between a verb and its direct object, correlated with reading ability: -0.3421, a significant (at $p<0.05$ level) and moderate correlation. Readers of both lower- and higher- proficiency are learning not to pause within an infinitive phrase or within a prepositional phrase to the same degree. However, the ability to recognize the relationship between a verb and its direct object, especially when the direct object is an infinitive phrase, which occurred once in the passage, or a noun clause, which occurred four times, and not pause inappropriately between them is accomplished much more often by the good readers, although the numbers of instances of these cases are too small to draw a definitive conclusion.

Conclusion

As a result of this study, two areas have emerged for further research. One area would call for research on the relationship between reading with correct thought groups and *speaking* spontaneously (rather than reading aloud) with correct breath groups (pausing). This relationship is difficult to investigate because "correct" pausing is much less well-defined in spontaneous speech. A phenomenon that native speakers use to compensate for pausing at inappropriate breaks is phrase repetition, that is, repeating from the beginning the phrase (breath group) in which the speaker paused inappropriately due, say, to the speaker's need to get his/her thoughts together.

Another area would focus on the correlation of the specific phrases with lower reading proficiency, for example, the prepositional phrase or the verb phrase with the direct object, could be taught explicitly as breath groups and then speech could be elicited from the subjects or the subjects could read a passage aloud to determine if their pausing before and within such phrases had improved and, if so, whether or not this improvement correlated with a higher reading score.

These are complex language issues, ones involving thought and language skills such as reading and speaking. Specific teaching/learning approaches can be introduced to facilitate the measurement of the variables. Drawing upon Palincsar, David, Winn and Stevens' (1991) notion of reciprocal teaching, in which students are taught strategies for reading, including generating questions, summarizing, clarifying, and predicting, second language learners can be taught the strategy of

1038

mouthing the words as they read in order to help them emphasize the one-to-one link between reading in thought groups and speaking in breath groups. Correct pausing, then, would transfer to both their reading skills and their speaking skills.

Students' proficiency in reading and in speaking could be measured using standard tests such as the TOEFL and the Test of Spoken English (TSE). From these could be inferred an ability, say, to read in thought groups better than speaking in breath groups. This observation has been noted, for example, with Asian students who usually score higher on the reading portion of the TOEFL than they do on the listening portion or on a speaking test. The cause-and-effect relationship could be inferred, treatment could be given, and the effect measured.

Viewing language and second language learning holistically, taking into account thought, reading (receptive language), and speaking (productive language), and researching these relationships can provide insight into second language learning.

References

Acton, W. R., Gilbert, J., & Wong, R. (1991, March). *Integrating pronunciation into the second language curriculum*. Paper presented at the 25th Annual International TESOL Convention, New York, NY.

Alderson, C. J., & Urquhart, A. H. (1984). *Reading in a foreign language*. New York: Longman.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*(4), 529-555.

Beebe, L. (1987). Myths about interlanguage phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 165- 175). Cambridge, MA: Newbury House.

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, & classroom perspectives*. Norwood, NJ: Ablex.

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford. CA: Stanford University Press.

Broselow, E. (1987). Non-obvious transfer: on predicting epenthesis errors. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 292-304). Cambridge, MA: Newbury House.

1039

Browne, S. C., & Huckin, T. N. (1987). Pronunciation tutorials for nonnative technical professionals: Aprogram description. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 45-57). Washington, DC: Teachers of English to Speakers of Other Languages.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(4), 561-584.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11, 51-62.*

Delamere, T. (1986). *The role of stereotyping in native speaker judgments of English as a second language learners' performance*. Unpublished dissertation, The Florida State University, Tallahassee, FL.

Dowhower, S. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice, 30*(3), 165-175.

Flege, J. E. (1981). The phonological basis of foreign accent: a hypothesis. *TESOL Quarterly, 15*, 443-455.

Flippo, R. (1984, July). *Evidence of the cognitive and metacognitive effects of punctuation and intonation: can the new technologies help?* Paper presented at the annual meeting of the United Kingdom Reading Association, Dundee, Scotland.

Gutknecht, C. (1978). Intonation and language learning: The necessity for an integrative approach. In A. Brown (Ed.), *Teaching English pronunciation* (pp. 259-275). London: Routledge.

Hatch, E. M. (1983). *Psycholinguistics: A second language perspective*. Rowel, MA: Newbury House.

Hieke, A. E. (1988). *Spoken language phonotactics*. Washington, DC: Center for Applied Linguistics. (ERIC Document Reproduction Service No. ED 290 320).

Johansson, S. (1978). *Studies in error gravity. Gothenburg studies in English: 44*. Goteborg, Sweden: Acta Universitatis Gothenburgensis.

Johnson, R. (1990). *The relationship between phonetic variables and holistic pronunciation ratings*. Unpublished Master's thesis, Iowa State University, Ames.

Johnson, R. (1993). *The relationship between native listener's perceptions of personality traits of nonnative speakers and the grammatical errors and phonetic errors in nonnative speakers' speech*. Unpublished doctoral dissertation, The Florida State University, Tallahassee.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Levelt, W. J. M. (1993). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

Ludwig, J. (1982). Native-speaker judgments of second-language learners' efforts at communication: A review. *Modern Language Journal, 66*, 274-283.

Maynard, S. (1986). *Variability in conversation management: Fragmentation of discourse and back-channel expressions in Japanese and English*. Washington, DC: Center for Applied Linguistics. (ERIC Document Reproduction Service No. ED 283 407).

Miccinati, J. (1985). Using prosodic cues to teach oral reading fluency. *The Reading Teacher, 39*, 206-212.

Pennington, M., & Richards, J. (1986). Pronunciation revisited. *TESOL Quarterly, 20*, 207-225.

Peregoy, S. (1989). Relationships between second language oral proficiency and reading comprehension of bilingual fifth grade students. *NABE Journal, 13*(3), 217-234.

Palincsar, A. M., David, Y. M., Winn, J. A., & Stevens, D. D. (1991). Examining the context of strategy instruction. *Remedial and Special Education, 12*(3), 43-53.

Radford, A. (1981). *Transformational syntax: A student's guide to Chomsky's Extended Standard Theory*. New York: Cambridge University Press.

Sato, C. J. (1987). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 248-260). Cambridge, MA: Newbury House.

Scanlan, T. (1987). Improving fluency in spoken French through a study of native pause behavior. *Foreign Language Annals, 20*(4), 345-352.

Tarone, E. E. (1987a). The phonology of interlanguage. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 70-85). Cambridge, MA: Newbury House.

Tarone, E. E. (1987b). Some influences on the syllable structure of interlanguage. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: the acquisition of a second language sound system* (pp. 232-247). Cambridge, MA: Newbury House.

1041

Taylor, D. S. (1981). Non-native speakers and the rhythm of English. In A. Brown (Ed.), *Teaching English pronunciation* (pp. 235-244). London: Routledge.

Wang, J. (1994). *The relationship between native English speakers', native Chinese speakers', and native Spanish speakers' attitudes toward English spoken by speakers from these three language backgrounds*. Unpublished Master's thesis, Southern Illinois University, Carbondale.

Wennerstrom, A. (1994). Intonational meaning in English discourse: a study of non-native speakers. *Applied Linguistics*, 15(4), 399-420.

Wong, R. (1987). *Teaching pronunciation: focus on rhythm*. Englewood Cliffs, NJ: Prentice-Hall.

Acknowledgments

The authors wish to thank Paul Angelis for his comments on earlier versions of this paper, the two anonymous reviewers for *Applied Language Learning* for their excellent suggestions for revision, Joseph Liberto for his advice on statistical tests, John Wendling for his assistance in tape-recording the oral reading samples, and Educational Testing Service for its permission to use the reading section of TOEFL in this research.

Authors

RUTH JOHNSON, Assistant Professor, Department of Linguistics, Southern Illinois University, Carbondale IL 62901-4517. *Specializations*: pronunciation (suprasegmentals), cross-cultural awareness, TESOL teacher training.

RITA MOORE, Instructor, Center for English as a Second Language, Southern Illinois University, Carbondale IL 62901-4518. *Specializations*: pronunciation (suprasegmentals), reading, cross-cultural awareness.

1042

Applied Language Learning
1997, Vol. 8, No. 1, pp. 43-73

A Gender-Related Analysis of Strategies Used to Process Written Input in the Native Language and a Foreign Language

Dolly Jesusita Young
University of Tennessee

Rebecca Oxford
University of Alabama

In this study, we examined the strategies used by native-English-speaking foreign language (FL) learners to read two Spanish texts and one English text. Our primary purpose was to investigate differences in FL reading strategies between males and females. We also posed the following questions: Were there significant gender differences in reading recall scores? Were there significant gender differences in self-reported levels of understanding and topic familiarity?

Forty-nine learners (26 females and 23 males) from a large southern university participated. After reading every passage, they conducted a think-aloud to report strategies they used. Think-aloud protocols were coded according to two strategy types: local and global. After this, subjects recalled the passage, that is, retold the story.

Results suggest that learners (whatever their level of language learning) processed the Spanish passage using primarily local strategies and the English passage using primarily global strategies. Results also indicate that males and females generally used similar strategies to process these passages, although there were specific strategies that were particular to males and females for each text type. There was also no significant difference in recall scores based on gender, although there were significant differences in recall scores based on the text type. Notably, though significant gender differences in self-reported levels of understanding and topic familiarity did not occur, there

© 1997, Dolly Jesusita Young and Rebecca Oxford

were significant differences among these variables according to text type. We offered a discussion of these results as related to previous gender-based research in strategy use. Finally, we provided implications for future work.

While a wealth of research exists on differences in language between males and females (Swann, 1992), the bulk of it examines language behavior as outwardly manifested in their speech. Only a limited amount of research has been conducted to determine whether there are differences in the way males and females process *input*, particularly in the FL context, and these few studies have produced somewhat conflicting results (Oxford, 1993, 1995). In addition, most of the research on input in language acquisition, irrespective of gender differences, has focused on auditory input (listening comprehension). This study examines whether differences exist in the strategies males and females use in processing written input in the native language (NL) and the FL.

Literature Review

NL Reading Strategy Research

A strategy is defined as a cognitive, metacognitive, social, or affective action plan used by a learner to learn material. Research in NL reading strategies suggests that poor readers tend to focus on reading as a decoding process, rather than a meaning-getting process (Baker & Brown, 1984). Golinkoff (1975-1976) offered a thorough comparison of reading comprehension processes of good and poor readers from first grade to college. According to Golinkoff, good comprehenders: (a) recognize words rapidly and accurately, (b) decode automatically, (c) read in phrase-like units, (d) are flexible in patterns of reading, (e) vary their eye movements, (f) shift the size of their processing units, (g) use supplementary contextual information efficiently, (h) pay attention to information relevant to their purpose and ignore irrelevant information, (i) read in the largest unit appropriate to the task, and (j) process the least amount of information compatible with the task (pp. 652-653). Poor NL comprehenders, by contrast, (a) are less able to organize texts, (b) decode texts slowly in a word-by-word manner with a minimum of text organization, and (c) are inflexible to variations in task demands (pp. 664-665). Golinkoff concluded that poor NL reading comprehension seems to be characterized by "being somewhat of a slave to the actual printed word and failing to extract structure and organization from a

text" (Golinkoff, 1975-76, p. 652). Thus, according to Golinkoff, poor NL readers laboriously and inflexibly use local strategies, rather than a combination of local and global strategies or a preponderance of global strategies.

FL Reading Strategy Research

The research on strategy use has also explored the basis for more successful and less successful reading comprehension in the FL or second language (L2) (e.g., Block, 1992; Hosenfeld, 1976; Munby, 1979; Pritchard, 1990; Sarig, 1987). This research, like research in the NL, frequently suggests that FL readers who approach reading as a "meaning-getting process" tend to be more successful than those who approach it as a "decoding process" (Carrell, 1992). With the exception of Pritchard and Sarig, however, the empirical data gathered in this research have been based on "small numbers of individual learners using think-aloud techniques" (Carrell, 1992, p. 121). Results often conflict (see Table 1).

Wolff argued that "L2 learners who have just begun L2 acquisition will almost exclusively resort to top-down [global] processing when reading texts or trying to decode utterances" (p. 313). Furthermore, Wolff contended that "the more difficulties L2 comprehenders have in decoding the incoming stimuli, the more productive they will become in activating concept-driven schemata [top-down, global strategies]" (p. 313).

Wolff examined processing strategies of German learners of English. Using two English language stories, *Rupert the Bear* and *The Balloon Story*, he compared students' (a) accurate text propositions, (b) propositions unrelated to the text, and (c) inferences. The latter story was more difficult with regard to linguistic aspects, content, and structure. Wolff used product-based data (recall protocols) as opposed to process-oriented data (think-alouds) in this study.[1] Wolff found a significant correlation between text difficulty and text content recall. Content from the easy story was recalled in more detail. For the difficult text, however, nonrelated propositions increased. Wolff concluded:

> Clearly, this difference points to the different processing strategies. The informant exposed to the more difficult text, whose bottom-up [local] processing was impeded by language deficiencies, used top-down [global] strategies instead. The high amount of inferences and nonrelated propositions in the *Balloon Story* protocols is a clear indication of an increased use of top-down strategies. (Wolff, 1987, p. 316)

TABLE 1
Results of NL and FL/ESL Reading Strategy Research

Study	NL Learners	FL or ESL Learners	Results
Baker and Brown (1984)	X		NL readers focused on local decoding, not meaning
Golinkoff (1975-1976)	X		Poor NL readers focus on local decoding inflexibly and laboriously
Wolff (1987)		X[1]	Novice FL readers used global strategies, especially when encountering difficult material
Hammadou (1991)		X[2]	Novice FL readers used global strategies, especially when encountering difficult material
Carrell (1983)	X	X	NL readers efficiently used global strategies; advanced high intermediate ESL readers often used local strategies, but not efficiently
Carrell (1989)		X	High proficiency ESL readers used global strategies efficiently; low proficiency Spanish readers used local strategies
Sarig (1987)	X	X	NL and FL readers processed texts the same way; global strategies led to both successful and unsuccessful NL/FL reading comprehension, but emphasizing local strategies led to unsuccessful NL/FL reading comprehension

[1]German-speaking learners of English
[2]English-speaking learners of Italian and French

1046

Hammadou (1991) found similar kinds of results in her examination of the relationships among content knowledge, inferencing strategies, and FL (Italian and French) reading comprehension of beginning and advanced (fourth-semester) learners. She discovered that the beginning readers' recall protocols tended to be longer than those of the advanced readers and that they contained more (global) inferences. Recall of the Italian passages had more illogical inferences than the French recalls. Hammadou explained that this difference might be due to the greater difficulty of the Italian texts. For both groups of students, linguistic proficiency was the significant factor and not content knowledge, thus offering support for Wolff's assumptions that novice learners rely primarily on concept-driven (top-down, global) processing when reading texts and that the more problems novices have, they more likely will resort to such top-down strategies.

Carrell (1983) examined several components of content knowledge as related to reading comprehension in the first and second language. Her subjects were native and non-native speakers of English. She found differences in processes between native English speakers and learners of English as a second language (ESL). The native speakers of English tended to use a top-down, global processing mode efficiently, while advanced and high intermediate ESL readers were neither efficient top-down nor bottom-up processors. These ESL readers seemed to be linguistically bound to the text, processing the "literal language of the text, but not making the necessary connections between the text and the appropriate background information" (p. 200).[2] Carrell (1988b) suggested that ESL readers seemed to engage primarily in text-bound or bottom-up (local) processes to the detriment of comprehension. She argued that "the most obvious cause of overreliance on the text in comprehension is the absence of relevant knowledge structures to utilize in top-down processing; if the schemata do not exist for the reader, they cannot be used" (p. 103).

In a later study, Carrell (1989) investigated ESL learners' perceptions of their metacognitive awareness regarding reading strategies. Using self-report questionnaires, she found that "the ESL group, at more advanced proficiency levels, tended to be more 'global' or top-down in their perceptions of effective and difficulty-causing reading strategies. The Spanish-as-a-foreign-language group, at lower proficiency levels, tended to be more local or bottom-up in their perceptions of effective and difficulty-causing reading strategies" (p. 128). She posited that language learners with lower proficiency levels might depend on local, word-oriented strategies when they read.

Yet Sarig (1987) suggested that readers do in fact process texts in the NL similarly to the way they process them in the FL. She analyzed readers' think-aloud protocols and found that the same strategies appeared to characterize processes in both languages. Top-down, global strategies led to both successful and unsuccessful reading comprehension. According to Sarig, an emphasis on more language-dependent, bottom-up, local strategies tended to impede comprehension in both NL and FL (p. 113).

Comparisons across these studies are problematic.[3] These studies, however, illustrate rather clearly the complex, sometimes conflicting nature of strategy use as related to comprehension.

General Gender-Based Research in the Foreign or the Second Language

As shown in Table 2, several studies found no differences between males and females in (a) foreign language skills (Feyten, 1991), (b) level of listening comprehension (Bacon, 1993; Markham, 1988[4]) and (c) rote memorization (Nyikos, 1990). However, other studies (see Table 2) found significant differences between males and females on the basis of various language-related factors.

For example, Boyle (1987) found male Chinese learners of English in China were stronger on vocabulary recognition in a listening task than females. Females, however, were stronger in overall language ability.

In two studies females surpassed males in ESL listening tasks. Farhady (1982) reported that females outperformed males in listening comprehension, and Eisenstein (1982) discovered that the former could discriminate among dialects better than the latter. Females were, in addition, better able to recognize what constituted more or lesser prestigious dialects.

Ludwig (1983) examined attitudes of French, German, and Spanish language learners toward individual languages and found that females considered FL listening tasks to be easier than did the males. On the other hand, males, compared with females, indicated FL speaking tasks were easier.

Research on Gender Differences in FL Learning Strategies

As displayed in Table 3 regarding research on gender differences in strategy use, females had a greater use of language learning strategies than males in most studies (see also Oxford, Nyikos, & Ehrman, 1988).

1048

TABLE 2
Gender Differences in Language Learning

Study	Languages Learned	No Gender Differences In...	Gender Differences In...
Feyten (1991)	Spanish	General FL skills	
Bacon (1993)	Spanish	Listening comprehension ability	
Markham (1988)	ESL	Listening comprehension ability	
Nyikos (1990)	German	Rote memorization ability	
Boyle (1987)	English in China		Aural vocabulary recognition (M>F); overall language ability (F>M)
Farhady (1982)	ESL		Listening comprehension (F>M)
Eisenstein (1982)	ESL		Aural discrimination of of dialects and prestige of dialects (F>M)
Ludwig (1983)	French, German, Spanish		Perception of ease of FL listening (F>M); perception of ease of FL speaking M>F)

For example, Politzer (1983) examined language learning behavior and social behavior and found that female FL learners used more social strategies than their male counterparts. In his study, females expressed interest in gentler aspects of interpersonal relationships such as cooperativeness and less interest in competitiveness and aggression. Ehrman and Oxford (1989) discovered that FL learners at the Foreign Service Institute showed many gender differences in strategy use, with females surpassing males in the use of general study strategies, strategies for meaning, self-management strategies, and functional practice strategies.

Oxford and Nyikos (1989) in a study of five FL's found that females used formal rule-based strategies, general study strategies, and conversational input-elicitational strategies more often than males. Nyikos (1987) examined the use of associative strategies for learning German noun clusters after strategy training and found significant differences between males and females.

TABLE 3
Gender Differences in Strategy Use of Native Speakers of English Learning Other Languages

Study	Languages Learned	Strategies Females Used More Than Males	Strategies Males Used More Than Females
Politzer (1983)	French, German, Spanish	Social strategies	None
Ehrman & Oxford (1989)	Many	General study strategies Strategies for meaning Self-management strategies Functional practice strategies	None
Oxford & Nyikos (1989)	French German Italian Russian Spanish	Formal rule-based strategies General study strategies Conversational input-elicitation strategies	None
Nyikos (1987)	German German	After training: Color-only memory strategies	After training: Color-plus-picture memory strategies
Oxford, Park-Oh, Ito, & Sumrall (1993)	Japanese	Cognitive strategies Social strategies Affective strategies	None
Wildner-Bassett (1992)	German Russian French	Compensation strategies Social strategies	None
Bacon & Finemann (1990)	Spanish	Global listening strategies Making friends with Spanish speaker (social strategy)	Local listening strategies

1050

TABLE 3 (continued)

Study	Languages Learned	Strategies Females Used More Than Males	Strategies Males Used More Than Females
Bacon (1993)	Spanish	Metacognitive strategies Cognitive strategies; (used formulaically)	Linear processing strategies; reference to native language
Brecht, Davidson & Ginsberg (1990)	Russian (in Russia)	None	Social strategies Affective strategies
Zoubir-Shaw & Oxford (1995)	French	Strategy types: Learning conjugations Learning grammar Learning from content Specific strategies: Using color-coded cards for gender; using pink/blue for gender; using other colors for gender; using flash cards; using organized lists; accepting rules at face value; reviewing textbook material	Strategy types: Learning from various activities Specific strategies: Concentrating more on oral communication than structures; being impeded by not knowing the meaning of a word (neg.); not comparing and accepting rules as a separate system (neg.); looking for the general meaning, idea or theme
Oxford et al. (1996)	Spanish	Strategy types: Cognitive strategies Memory strategies Social strategies Affective strategies Specific strategies: Trying out new vocabulary learning strategies	Strategy types: None Specific strategies: Thinking about my progress: judging success of a given strategy

In a study of high school students learning Japanese, Oxford, Park-Oh, Ito, and Sumrall (1993) found that females tended to use cognitive strategies, social strategies, and affective strategies more often than males. Wildner-Bassett (1992) examined gender as related to FL learning strategies and found that unlike males, females reported "using more strategies for coping with missing knowledge in their language learning [compensation strategies] and generally preferred to learn with others [social strategies]" (p. 6). Bacon and Finemann (1990) also found a social

interaction variable in strategy use of men and women learning Spanish. In their study, women significantly more often than men expressed that they had friends who were Spanish speakers. In this same study, Bacon and Finemann found significant differences in learning and comprehension strategies in the self-reports of males and females. "Female subjects were more likely to report using global/synthetic strategies, whereas males were significantly more likely to report using [local] decoding/analytic strategies" (p. 468). Bacon and Finemann used the distinction between global and local strategies that we employ in our present study.

In another study on FL strategy use and gender among learners of Spanish, Bacon (1993) reported no significant difference in the listening comprehension levels of males and females, but did find some interesting results regarding strategy use. Those differences depended on the difficulty level of the passage and the order in which they heard the passage.

> Women were more apt to report using metacognitive strategies, and were more likely to adjust that usage when passage difficulty demanded it. Women used cognitive strategies (both bottom-up and top-down) almost formulaically, as if they had rehearsed their listening procedures, and always applied the same outline. Men dealt with the difficult passage more aggressively, with strategies that depended on linear processing and reference to English. (pp. 171-172)

Zoubir-Shaw and Oxford (1995) studied gender differences in university students of French. They discovered that females used a number of general strategy categories (related to learning conjugations, learning grammar rules/structures, and learning from context) significantly more often than males. Females also employed a number of specific strategies more often than males: using color coding, flash cards, organized lists; accepting rules, and reviewing textbooks. On the other hand, males more often than females used the general strategy category of learning from various activities, and males also used certain specific strategies (especially "negative" ones) more often than females.

Oxford, Lavine, Hollaway, Felkins, and Saleh (1996) found gender differences in learning strategy use via diaries of university-level Spanish language students. Several of the general strategy-category differences favored females, and none favored males. For instance, significantly more females than males used cognitive, memory, social, and

1052

affective strategies. In terms of specific strategies rather than overall categories, the strategy of trying out new techniques for vocabulary learning was used significantly more often by females than males. However, some gender differences favoring males were seen for two specific evaluation-related strategies (thinking about progress and judging success of a given strategy).

Interestingly, in the only research conducted in informal language development of American students in the target country, Brecht, Davidson, and Ginsberg (1990, 1991) discovered some of the expected gender differences to be reversed. American male college students in study-abroad programs in Russia (learning Russian as an L2) were more likely to improve their proficiency, more likely to use social and affective strategies, and more likely to employ a broader range of strategies than females. Brecht et al. (1990, 1991) at first suggested that the greater aggression of male college students allowed them greater access to the foreign culture and to strategy-using opportunities, but these investigators subsequently found that males had more opportunity because of the male-oriented society in Russia.

In addition to these FL studies of native English speakers learning an additional language, a whole set of studies exists involving English as a second or foreign language, revealing gender differences in language learning strategy use (see detailed reviews in Oxford, 1993, 1995; Zoubir-Shaw & Oxford, 1995). These studies typically followed along the same lines as most of the studies of native English speakers learning an additional language, with females using strategies significantly more often than males in most investigations.

Oxford (1993) summarized the gender-based research in learning strategies in the following way:

> Whenever strategy research has considered gender, it has usually demonstrated gender differences in strategy frequency, with females choosing to use particular sets of strategies more often than males. Females especially tended to use general study strategies, social strategies, affective strategies, and certain conversational or functional practice strategies more frequently than males across a number of studies, usually showing a greater range of frequently used strategy categories. (p. 83)

To date, no study has examined strategies used by males and females in processing written input in both their NL and their FL. This study is an attempt to begin to fill that void in the hope of contributing to gender-based research in FL learning.

Method

Research Questions

1. Are there significant differences in the types of strategies learners use to process texts in the FL and the NL?
2. Are there significant differences in the types of strategies male learners and female learners use to process FL and NL texts?
3. Are there significant differences in FL and NL recall scores?
4. Are there significant differences in FL and NL recall scores between males and females?
5. Are there significant differences in self-reported levels of understanding and familiarity with passage topics between males and females?

Participants

A total of 49 native-English-speaking students (26 females and 23 males) participated in the study: 14 first-year, 14 second-year, 9 third-year, and 12 fourth-year students of Spanish. Students in the first year were in their second semester of Spanish, and second-year students were in their fourth semester. Students in the third-year courses were in their sixth semester of Spanish. One of the third-year courses was a conversation class, and the other a literature reading class. Students in the fourth-year courses had completed at least seven semesters of Spanish or the equivalent. One of the senior-level courses was in Spanish civilization and the other in applied linguistics. They all participated in three separate think-aloud procedures, one in English and two in Spanish, and a corresponding reading recall task for each passage.

Materials

Students read three different kinds of passages. Two were in Spanish, and one was in their native language, English. Three samples of Spanish-edited passages were taken from textbooks used at the course levels of the students.[5] For first- and second-year students, passages were selected from chapters they had not yet read in their Spanish textbook. The passage for first-year students was 388 words in length, and the second-year passage was 391 words, including the title and captions. One passage of 481 words was selected for third- and fourth-year students. This passage was taken from the Spanish civilization textbook, which was used by both third- and fourth-year students. All three of these passages had been written and edited for students at particular levels of language learning, and all three had cultural themes.[6] None

1054

had been previously read by the participants. These passages were linguistically and culturally challenging. They included newly-acquired vocabulary and newly-learned structures. In addition, students had little background knowledge about the topics of these passages.

The topic of the first-year Spanish-edited passage was Hispanic-American economics of the past and the future. The topic for the second-year Spanish-edited passage was the presence of foreign cultures in work and leisure of the Hispanic world. The topic for the other Spanish-edited passage dealt with Hispanic historical themes.

The second Spanish passage, completely authentic, was taken from a popular Spanish magazine similar to *Good Housekeeping*, and it focused on myths and American medical findings about chocolate. This passage was linguistically and conceptually accessible for students because of its (a) content familiarity, (b) use of visual cues, and (c) consistent use of contrast in its rhetorical organization.

The English passage, taken from *Scientific American*, was about the scanning tunneling microscope. This authentic text was selected because of its challenging content. Students did not have much knowledge about this topic, thus they had to use strategies in their native language.

Instruments

To determine content knowledge, students rated their familiarity with the passage topic on a scale from 1 (very familiar) to 3 (not familiar). To account for self-assessment of their comprehension of the passages, students rated their understanding of the passage on a scale from 1 to 5 (1 = all of it, 2 = most of it, 3 = about half of it, 4 = some of it, 5 = none of it).

Think-aloud protocols assessed learners' reading strategies. These protocols were coded using standard classifications employed in previous studies (Anderson, 1991; Block, 1986; Carrell, 1989; Kletzien, 1991; Pritchard, 1990). The classification scheme used in this study divided strategies into two rubrics. The first rubric contained strategies which focused on reading as a decoding process (i.e., those strategies pertaining to "sound-letter, word-meaning, sentence syntax, and text details") (Carrell, 1989, p. 126) such as skipping specific unknown words, breaking lexical items into parts, translating a word or a phrase, and paraphrasing. The second rubric contained strategies which focused on reading as a meaning-getting process (strategies pertaining to "background knowledge, text gist, and textual organization") (Carrell, 1989, p. 126), such as anticipating content, integrating information, recognizing text structure, and using background knowledge (see Table 4 for strategy classifications).

TABLE 4
Strategy Classification Scheme

Strategy	Definition	Sample Responses
Local Strategies:		
States understanding of words/vocabulary.	The reader acknowledges comprehension based on knowing all the words.	"Oh, this was easy to read 'cause the vocabulary was easy. I didn't have a problem. I seemed to know most of the words."
Skips specific unknown words.	The readers states that he/she skipped a word that was not known.	"I just kept on reading and if I didn't know a word I did not stop, I just skipped it."
Expresses use of gloss.	The reader voices use of word glosses or a need for a gloss.	"I read the gloss for this word because I didn't know what it meant."
Breaks lexical items into parts.	The reader breaks up words or phrases into smaller units to figure out the word/phase.	"Meaningless. Mean is *significado* but less is *bajo significado.*"
Uses cognates between NL and FL to comprehend	The reader expresses ease of understanding because of words that look and mean the same in NL/FL.	"*Conservar* was easy 'cause it looks like what it is in English."
Solves vocabulary problems.	The reader uses context, a synonym, or some other word-solving behavior to understand a particular word.	"Straight-forward grammar, means easy grammar."
Translates a word or phrase into FL.	The reader expresses meaning of word or phase or FL.	"I just put the words in Spanish."
Questions meaning of a word.	The reader does not understand a particular word.	"I don't understand this word."
Identifies, through circling, underlining, or placing an arrow, words/phases not understood.	The reader states that he/she circled, underlined, etc., a word or phase not understood.	"I circled words I didn't know and went back to figure them out if I could."

1056

TABLE 4 (continued)

Strategy	Definition	Sample Responses
Questions meaning of a clause or sentence.	The reader indicates that he/she does not understand the meaning of a portion of the text.	"What's this sentence mean?"
Uses knowledge of syntax and punctuation or other grammar.	The reader expresses awareness of grammar, syntax, and parts of speech or punctuation.	"I put *taking* because I knew it had to be a verb. I figure x because there was a comma there. The word order here, Spanish is kind of reverse order."
Monitors reading pace and reading behavior.	The reader makes reference to slowing down, rereading, or perhaps reading on in spite of not understanding some things. The reader mentions specifically that he/she went back and read something again, or the reader indicates using information which is more than a sentence away.	"I just slowed down if I didn't something." "Even though I wasn't getting everything, I just kept reading."
Paraphrases.	The reader rewords the original wording of the text.	"Reading through the first paragraph, talking about the smallness of the world and how international commerce and tourism and media and the arts show other cultures are around."
Global Strategies:		
Skims, reads headings, subtitles; looks at pictures.	The reader previews text to get a general idea of what the article is about before actually reading the text.	"Well, I just looked it over quickly before I read it to get an idea of what it was going to be."
Anticipates content.	The reader predicts what content will occur in succeeding portions of text.	"I guess the story will be about how you go about talking to babies."

1057

TABLE 4 (continued)

Strategy	Definition	Sample Responses
Recognizes text structure.	The reader distinguishes between main points and supporting details or discusses the purpose of information or notes how the information is presented.	"This article just compares the myths and realities of chocolate."
Integrates information.	This reader connects new information with previously stated content.	"Oh, this connected with the sentence just before it."
Reacts to the text.	The reader reacts emotionally to information in the text.	"I love chocolate." "It was all pretty easy."
Speculates beyond the information in the text.	The reader shares a thought that goes beyond the information contained in the text.	"I was thinking about my roommate who loves chocolate. She needs to read this."
Acknowledges lack of background knowledge.	The reader states lack of familiarity with or knowledge about text topic.	"I just don't know much about chemistry, biology, etc., so this was hard to understand."
Reads ahead.	The reader specifically mentions reading ahead as he/she reads.	"I looked at the next subtitle *cholesterol y cafeina* and got the idea that that's where they were ending up."
Visualizes.	The reader indicates that he/she had a mental image.	"I picture the needle like an airplane going over ridges."
Identifies main idea.	The reader related major points of paragraph or passage.	"This whole thing was talking about how Africa was trying to get independence so they have to have control of the government."

1058

TABLE 4 (continued)

Strategy	Definition	Sample Responses
Uses inference or draws conclusions.	The reader indicates that he/she guesses based on information in text and own knowledge.	"I wasn't familiar with either of these names, so I simply used the fact that Charles Arden-Clarke was not African, and Nkrumah, who was the Gold Coast Prime Minister, he was getting advice, so I would assume that Clarke was an advisor. I knew he wasn't an African because of his name."
Uses background knowledge.	The reader states a familiarity or knowledge about text topic.	"I just thought about the things I learned in Food and Nutrition." "I thought about what I experienced from *Pasajes*."

In scoring the think-alouds, interrater reliability was established by the random recoding of a 20% sample of the tapescripts by the first author and an independent rater. Interrater reliability was assessed for all three texts (Spanish-authentic, Spanish-edited, and English passages). Thus, for all three passages, coefficients of .80 to .85 were achieved for the number of strategies used. For all 25 strategy types, interater reliability ranged from moderate to strong.

Recall tasks assessed reading comprehension. This investigation used Bernhardt's (1988) procedure for the development and scoring of the recall protocols because of its high reliability and validity. While other recall procedures exist, such as the Meyers system, Bernhardt illustrated several advantages of the propositional and weighted scoring system used in this study. The first step in setting up the recall protocol scoring templates consisted of identifying pausal units. A pausal unit is a natural break or breath group under a normal oral reading of a text. Pausal units were identified by a total of six fluent readers of Spanish (three native and three non-native speakers). Whenever there were differences, the narrower units were selected as defined in Bernhardt (1988). Second, weights were assigned to the pausal units from 3 (most significant) to 1 (least significant) in terms of the unit's importance to the message of the passage. In cases of disagreement, a consensus determined the weight of the unit in question.

In scoring the recall protocols, interrater reliability was established by the random recoding of a 20% sample of the tapescripts by the researcher and an independent rater. Interrater reliability coefficients of

1059

.89 (for level 3 propositions), .81 (for level 2 propositions) and .58 (for level 1 propositions), $p < .01$, were achieved. The low reliability coefficient for level 1 propositions may be due to the few identifications of level 1 propositions (which are the least significant to text comprehension), thereby making any differences appear greater.

Procedure

Students read the two Spanish (edited and authentic) passages first, followed by the English passage. The order of the Spanish passages was varied to control for possible ordering effects. For each passage, student reading time was documented. Immediately after each reading, students recounted in their native language (English) everything they thought about during the reading that helped them understand what they were reading.[7] The resulting think-aloud protocols formed the data base for strategy analyses. Then students were asked to "recall everything you remember from this text" in English.[8] The resulting oral recall protocols served to assess reading comprehension. After all passage tasks had been completed, students rated the content familiarity they brought to the passage before reading it and their understanding of the passage. All data were recorded on tape and transcribed.

Analysis and Results

Findings by Research Questions

This study proposed several research questions that will now be addressed.

1. Are there significant differences in the types of strategies learners use to process texts in the FL and the NL? To address this question, the think-alouds for each passage were analyzed for number of local strategies (strategies involving "sound letter, word-meaning, sentence syntax, and text details," Carrell, 1989) as compared to global strategies (strategies pertaining to "background knowledge, text gist, and textual organization," Carrell, 1989). Students used almost twice as many strategies to process the FL Spanish passages (the edited passage mean was 15.61; the authentic passage mean was 13.08) than to process the NL English passage (mean was 7.78). Students used a similar number of global strategies for all three passages (see Figure 1). They used the most local strategies for the Spanish-edited passages, somewhat fewer for the FL Spanish-authentic passage, and even fewer for the NL English passage. A t-test measured the difference between means of local

1060

and global strategy use for each text type. Results indicated that there were significant differences in the mean number of global strategies versus local strategies used by the participants to process the FL Spanish-edited passages ($p < .001$), the FL Spanish-authentic passage on myths about chocolate ($p < .003$), and the NL English passage ($p < .004$). In other words, not only was there a significant difference in strategy use between text types in the NL versus the FL, there was also a significant difference in strategy use between the two FL text types (the Spanish-edited passage and the Spanish-authentic passage).

Further analysis of strategy-type frequency by level and text type indicated that only in second-year Spanish was strategy use significantly different for the Spanish-edited passage and for the Spanish-authentic passage. The difference in strategy-type frequency between these two text types was due to an extensive use of local strategies by students in second-year Spanish to process the Spanish-edited passage.

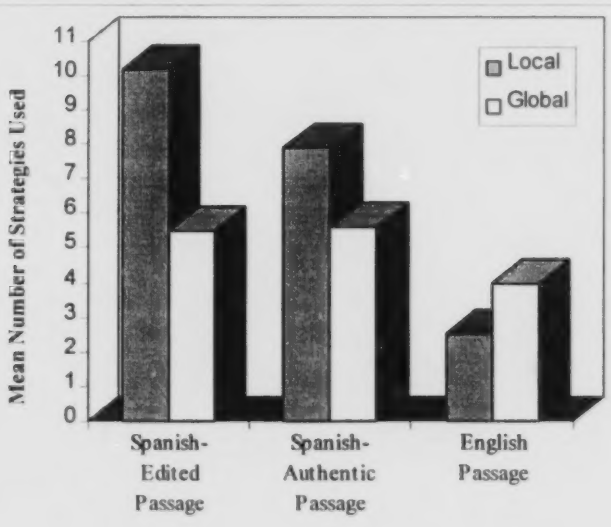


Figure 1. Local and global strategy use by passage (text type)

2. Are there any significant differences in the types of strategies male learners and female learners use to process FL and NL texts? Means are shown in Table 5. A t-test examined the difference between male and female mean frequencies of types of strategies for each text. There were no significant differences between the means of the types

1061

of strategies males and females used to process the passages in the FL or the NL. In other words, there was no difference in mean use of strategy-type (local versus global strategies) by gender.

However, through the use of chi-square tests, significant differences between males and females were found regarding the frequency of use of specific strategies. Males more often than females used the following specific strategies: strategy 12 - monitors reading pace and reading behavior (chi-square = 4.93, df = 1, $p < .026$, for the FL Spanish-edited passage); strategy 13 - paraphrases (chi-square = 5.52, df = 1, $p < .019$, for the FL Spanish-authentic passage); and strategy 1 - states understanding of words (chi-square = 4.93, df = 1, $p < .026$, for the NL English passage). Females more often than males employed the following specific strategies: strategy 6 - solves vocabulary problems (chi-square = 4.47, df = 1, $p < .035$, for the Spanish-authentic passage); and strategy 20 - acknowledges lack of background knowledge (chi-square = 5.85, df = 1, $p < .016$, for the NL English passage).

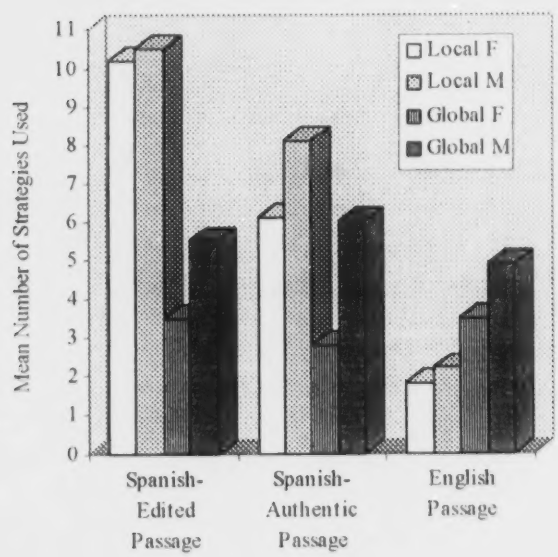


Figure 2. Local and global strategy use by passage (text type) by gender

1062

TABLE 5

Means of Local and Global Strategies by Text Type and by Gender

Gender	Strategy Type	Means for Spanish-Edited Passage	Means for Spanish-Authentic Passage	Means for English Passage
Females				
	Local	10.19	6.96	2.65
	Global	4.34	3.73	4.30
Males				
	Local	10.52	8.90	2.95
	Global	6.30	6.80	5.73

3. Are there significant differences in the FL and NL recall scores? Recall scores were 9 points higher for the FL Spanish-authentic passages (mean 24%) than the Spanish-edited passages (mean 15%).[9] The recall scores for the NL English passage were higher than for the Spanish-edited passages but lower than the FL Spanish-authentic passage (mean 16%). T-tests on recall scores for these passages indicated that there was a significant difference in recall scores for the Spanish-authentic and Spanish-edited passages ($t = 5.50$, $p < .0001$), indicating that students recalled more from the authentic passage than the edited passages, which had been purposely written and edited for language textbooks! There was no significant difference, however, in recall scores of the FL Spanish-edited passages and the NL English passage. Finally, there was a significant difference in the recall scores between the FL Spanish-authentic passage and the NL English passage ($t = 6.00$, $p < .0001$).

4. Are there significant differences in FL and NL recall scores between males and females? Results of a t-test of recall scores by gender indicated that there were no significant differences in recall scores by gender for any of the three text types.

5. Are there significant differences in self-reported levels of understanding and familiarity with passage topics between males and females? Results of t-tests indicated that there were no significant differences by gender in reported levels of understanding the passages or in ratings of

1063

familiarity with passage topics for any of the three passages. In other words, there was no difference between males and females in the background information (or lack thereof) the readers brought to each text. There was also no difference between males and females in how well they thought they understood each passage.

Discussion

In this study, the strategies learners used to process the NL (English) and FL (Spanish) passages were slightly different. Learners used significantly more local strategies to process the FL passages and predominantly global strategies to process the NL passage. Furthermore, regarding the two FL passages, learners used more local strategies to process the edited passages compared to the authentic ones. This suggests that the perceived difficulty of the text and whether the text is NL or FL might be related to the general kinds of strategies readers employ. Carrell (1983) showed that readers used global strategies efficiently with NL texts but that ESL learners were not consistently effective with local strategies. The use of global strategies with difficult FL or ESL material was shown by Wolff (1987) and Hammadou (1991) but not by Carrell (1989) and not in our present study. Sarig's (1987) finding that FL and NL readers processed reading passages the same way was obviously not replicated in our study.

When learners' use of strategy type (global vs. local) was examined by gender, there were no significant differences, though an interesting tendency occurred. Females tended to use global strategies slightly more often than males, though this result did not reach full significance. In Bacon and Finemann's (1990) study of strategies used for FL listening comprehension, females also more frequently than males used global strategies, and males more often than females used local strategies.

Previous research has sometimes found differences in strategy use as a function of the difficulty level of a passage. For example, in Bacon (1993), females but not males adjusted their strategies according to the difficulty level of the FL passages. In our study, students of both genders tended to use more local strategies with the more difficult FL Spanish text type, the edited cultural passages. Students in second-year Spanish used significantly more local strategies to process their Spanish-edited passage than students used at other instructional levels. This suggests that the Spanish-edited passages for the second year might have been even more difficult than the Spanish-edited passages for the first and third/four years. In any event, because both male and female strategy use, even for second-year Spanish, was not significantly different,

1064

strategy adjustments based on text difficulty in this study were not related to gender.

In looking at the specific strategies in the present study, there were some interesting gender-related results. We found gender differences in the use of specific strategies. Here is how these strategies related to each of the passages and to previous research on gender differences in strategy use:

a. With the *FL Spanish-edited passage*, males more often than females reported monitoring the reading pace and reading behavior (strategy 12). This is similar to the monitoring/evaluative behavior shown more often by males than females in the study of university Spanish students conducted by Oxford et al. (1996)

b. With the *FL Spanish-authentic passage*, females tried more often than males to solve vocabulary problems (strategy 6). This is completely in line with the Oxford et al. (1996) finding that female university learners of Spanish tried out new techniques for vocabulary learning the Spanish-authentic passage, men more frequently than women reported the use of paraphrasing (strategy 13). Similarly, Zoubir-Shaw and Oxford (1995) found that male university students of French looked for general meanings, ideas, and themes, a strategy that frequently involves or leads to paraphrasing.

c. With the *NL English passage*, some differences occurred by gender. Males significantly more often than females expressed understanding the words in this passage (strategy 1), and females more often than males acknowledged their lack of background knowledge about the passage topic (strategy 20), which was about scanning tunneling microscopes. Males might be more likely to be interested in this topic than females would; however, there was no expressed significant difference in topic familiarity with any of the three passages by gender. Overall, we might say that this particular passage topic was more male-friendly than female-friendly. Nevertheless, females understood just as much of the general meaning of the NL English passage as did the males, as the lack of any significant difference in recall scores suggests.

Earlier research has indicated the superiority of females over males in many aspects of language performance (Oxford, 1993, 1995). However, in the current study, females did not outperform males in the reading recall tasks. Both males and females performed similarly.

What do the results of the present investigation, taken together, suggest about FL reading strategies and gender-based differences in reading strategies? This research indicates that although there are indeed some contrasts in the way males and females process reading passages, these differences might not occur in overall, general categories of

1065

strategy use such as global versus local. The more meaningful contrasts might reside at the level of specific strategies, the identifiable behaviors that learners use to aid them in reading comprehension. Even at that highly specific strategy level, the similarities in reading strategies across males and females are greater in number than the differences.

Thus, there may be few inherently "male" or "female" reading strategies. Should we therefore ignore the gender differences that we do find? No, we definitely should not. There seem to be some recurring patterns in the ways that males and females go about learning a foreign language, at least at the level of specific strategies. The practical significance of these findings is that teachers need to watch for and understand such gender differences. For example, if future research shows that males need to improve specific strategies, such as solving vocabulary problems, teachers can provide this kind of assistance. If further studies show that females could benefit from developing strategies in self-monitoring and self-evaluation, teachers can help them as well. Moreover, teachers must try to help all learners, female and male, to use appropriate strategies. We need to help everyone develop optimal language learning strategies that go beyond gender boundaries. All students, male and female, can learn to compensate for many current weaknesses in strategy use and can develop a larger strategy repertoire, one that fits with the requirements of the language task and with the learning style and personality of the student.

Our study of gender differences raises additional issues for future research. First, investigators need to consider the differences in strategy issues that occur at various reading difficulty levels. The difficulty level of a text is determined by a number of variables and is often dependent on what the reader brings to a text, such as the reader's proficiency level, background knowledge, and strategy preferences. In a text, language complexity and text type can also help determine a text's level of difficulty. As FL reading researchers, we constantly ask how these variables interact and which ones seem to play the key roles in FL reading comprehension. In this study, for example, the FL Spanish-edited passages were more difficult for students because, unlike the FL Spanish-authentic passage, (a) the topics in these texts were unfamiliar to the learners, (b) these passages were more linguistically dense, and (c) they did not use obvious rhetorical formats to organize the content. Thus, students used many more local strategies than global to process these passages. To what extent did the rhetorical features of the FL Spanish-authentic passage enhance reading comprehension? If the FL Spanish-edited passages had been formatted like an authentic text, and if they had used obvious rhetorical structures to organize the content, would

1066

learners have used fewer local strategies than they did in this study? We need more research into what determines text difficulty and how text difficulty relates to specific-strategy frequency and strategy-type frequency.

Second, researchers should examine the relationship between strategies and reading competence. The use of certain strategies might be an indicator of reading competence. Third, investigations are needed concerning the amount of cognitive energy consumed by local versus global processing. If local strategies can become automated, readers might not need to make a conscious effort to use such strategies, thus freeing up cognitive resources for global or conceptually-driven processing. Finally, scientists should carefully consider different classifications of language learning strategies, particularly in the area of reading (global versus local, cognitive versus metacognitive versus affective/social). Classification schemes could be evaluated by their usefulness and the insights they yield depending upon the purpose for which they are used.

Conclusion

Research on gender differences in specific-strategy frequency and strategy-type frequency could offer insights into the similarities and differences between male and female cognitive processes in language acquisition. This research could also contribute to our understanding of the reading comprehension process. This study examined strategies of males and females in processing written input in both the learner's native language, English, and a foreign language, Spanish, in an attempt to expand gender-related research in language acquisition and contribute further data to research in FL reading. The general lack of significant gender differences in frequency of specific-strategy use and strategy-type (local versus global) use should spark queries by researchers who study gender differences in language acquisition. However, the observed differences in the use of some specific strategies by gender should also generate questions about those differences for future researchers and should emphasize the need to look at particular strategies rather than just at more general strategy types. Finally, the strategy differences in the NL and FL text types and instructional levels also contribute to our evolving understanding of processes in FL reading.

1067

Notes

[1] Recall protocols are transcribed descriptions of the contents of the passage, and think-alouds are oral descriptions of cognitive processes while reading.[8] Lee (1986) suggested that assessing comprehension in the target language does not give an accurate measure of reading proficiency because students can comprehend better than they can articulate.

[2] This study did not distinguish good readers from poor readers.

[3] See Bernhardt's chapter "The Data Base in Second Language Reading" in her book *Developing Reading* (1991).

[4] Interestingly, Markham (1988) found that EFL listeners attended to male speakers more than female speakers. Markham concluded that "gender bias is a pervasive factor that exerts an influence on ESL students' recall of orally presented material" (p. 404).

[5] Both first-year and second-year passages were from McGraw-Hill publications: T. Dorwick, M. Marks, M. Knorr, B. VanPatten, & T. Higgs (1987), *?Qué Tal?*, 2nd ed., New York: Random House, p. 399; and M. L. Bretz, T. Dvorak, & C. Kirschner (1987), *Pasajes*, 2nd ed., New York: Random House, pp. 235-236. The third passage was from V. Cantarino (1988), *Civilización y cultura de España*, 2nd ed., New York: Macmillan, pp. 109-110.

[6] Thalia Dorwick, Publisher, Foreign Language College Division, confirmed that the first-year and second-year passages were created for language learners who were at a particular level of instruction.

[7] The first author preferred retrospective versus concurrent think-alouds to avoid interruptions in the reading process.

[8] Lee (1986) suggests that assessing comprehension in the target language does not give an accurate measure of reading proficiency because students can comprehend better than they can articulate.

[9] Bernhardt and Everson (1988) reported an average recall score of 20% (30% for upper levels) for learners in their study. They pointed out that even a native speaker would not recall more than 80%. They cite 20% to 30% recall as quite remarkable for an edited text.

1068

References

Anderson, N. J. (1989). *Reading comprehension tests versus academic reading: What are second language readers doing?* Unpublished doctoral dissertation, University of Texas, Austin.

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal, 75,* 460-472.

Bacon, S. M. (1993). The relationship between gender, comprehension, processing strategies and cognitive and affective response in foreign language listening. *Modern Language Journal, 76,* 160-178.

Baker, L. & Brown, A. L. (1984). Metacognitive skills in reading. In P. D. Person (Ed.), *Handbook or research on reading* (pp. 353-394). New York: Longman.

Benedetto, R. A.. (1984). *A psycholinguistic investigation of the top-level organization strategies in first and second language reading: Five cases studies.* Unpublished doctoral dissertation, New York University.

Bernhardt, E. B. (1988). Merging research and assessment perspectives in foreign language reading. (Technical Report) Washington, DC Institute of Advanced Studies, Mellon Fellowship Program, National Foreign Language Center.

Bernhardt, E. B., & Everson, M. E. (1988). Second language reading: A cognitive perspective. In G. Luer & U. Laff (Ed.), *Proceeding of the Fourth European Conference on Eye Movements* (Vol. 1, pp. 1-44). Amersterdam: Elsevier Science Publishers.

Bernhardt, E. B. (1991). *Reading development in a second language.* Norwood, NJ: Ablex.

Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly.* 20(3), 463-491.

Block, E. (1992). See how they read: Comprehension monitoring of L1 and L2 readers. *TESOL Quarterly,* 26, 319-342.

Boyle, J. P. (1987). Sex differences in listening vocabulary. *Language 37,* 3273-3284. *Learning, 37,* 3273-3284.

Brecht, R., Davidson, D., & Ginsberg, R. (1990). The empirical study of proficiency gain in study abroad environments among American students of Russian. In D. Davidson (Ed.), *American contributions to the VII International Congress of MAPRIAL.* Washington, DC: American Council of Teachers of Russian.

1069

Brecht, R., Davidson, D., & Ginsberg, R. (1991). On evaluating language proficiency gain in study abroad environments: An empirical study of American students of Russian (A preliminary analysis of data). In Z. D. Dabars (Ed.), *Selected papers delivered at the NEH Symposium in Russian Language and Culture*. Baltimore. CORLAC/Friends School of Baltimore.

Carrell, P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning, 33*, 183-207.

Carrell, P. L. (1988). Schema theory and ESL reading: Classroom implications and applications. *Modern Language Journal, 68*, 332-343.

Carrell, P. L. (1989). Metacognitive awareness and second language reading. *Modern Language Journal, 73*, 121-134.

Carrell, P. L. (1992). Awareness of text structure: Effects on recall. *Language Learning, 42*, 1-20.

Clarke, M. A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning, 29*, 21-50.

Coady, J. (1979). A psycholinguistic model of the ESL reader. In R. Mackay, B. Barkman, & R. R. Jordan (Eds.), *Reading in a second language: Hypotheses, organization, and practice* (pp. 5-12). Rowley, MA: Newbury House.

Cummins, J. (1980). The cross-lingual dimension of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*, 175-187.

Cziko, G. A. (1980). Language competence and reading strategies: A comparison of first and second language oral reading errors. *Language Learning, 30*, 101-116.

Ehrman, M., & Oxford, R. (1988). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *Modern Language Journal, 72*, 253-265.

Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning, 32*, 367-391.

Farhady, H. (1982). Measures of language proficiency from the learner's perspective. *TESOL Quarterly, 16*, 43-59.

Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. *Modern Language Journal, 75*, 173-180.

Golinkoff, R.M. (1975-1976). A comparison of reading comprehension processess in good and poor comprehenders. *Reading Research Quarterly, 11*, 623-659.

1070

Hosenfeld, C. (1976). A preliminary investigation of the reading strategies of successful and unsuccessful second language learners. *System, 5,* 110-123.

Hammadou, J. (1991). Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *Modern Language Journal, 75,* 27-38.

Hudson, T. (1982). The effects of induced schemata on the "short-circuit" in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning, 32,* 1-31.

Kletzien, S. B. (1991). Good and poor high school comprehenders' use of reading strategies for expository texts of differing levels. *Reading Research Quarterly, 26,* 67-86.

Lee, J. F. (1986). On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition, 8,* 83-93

Ludwig, J. (1983). Attitudes and expectation: A profile of female and male students of college French, German, and Spanish. *Modern Language Journal, 67,* 216-227.

Macnamara, J. (1970). Comparative studies of reading and problem solving in two languages. *TESOL Quarterly, 4,* 107-116.

Markham, P. L. (1988). Gender differences and the perceived expertness of the speaker as factors in ESL listening recall. *TESOL Quarterly, 22,* 297-406.

Munby, J. (1979). Teaching intensive reading skills. In R. Mackay, B. Barkman, & R. R. Jordan (Eds.), *Reading in a second language: Hypotheses, organization, and practice* (pp. 142-158). Rowley, MA: Newbury House.

Nyikos, M. (1987). *The effect of color and imagery as mnemonic strategies on learning and retention of lexical items in German.* Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *Modern Language Journal, 74,* 273-287.

Oxford, R., Nyikos, M., & Ehrman, M. (1988). *Vive la différence?* Reflections on sex differences in use of language learning strategies. *Foreign Language Annals,* 21, 321-329.

Oxford, R. (1993). Instructional implications of gender differences in second/foreign language learning styles and strategies. *Applied Language Learning, 4,* 65-94.

Oxford, R. (1995). In J. Reid (Ed). *Learning styles in the ESL/EFL classroom.* Boston: Heinle & Heinle.

1071

Oxford, R. L., Lavine, R. Z., Hollaway, M. E., Felkins, G., and Saleh, A. (1996). Telling their stories: Language learners use diaries and recollective studies. In R. L. Oxford (Ed.). *Language learning strategies around the world: Crosscultural perspectives.* Manoa: University of Hawaii Press.

Oxford, R., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal, 7* 219-300.

Oxford, R., Park-Oh, Y., Ito, S., & Sumrall, M. (1993). Learning a language by satellite: What influences achievement? *System, 21* (1), 31-48.

Politzer, R.L. (1983). An exploratory study of self-reported language learning behaviors and their relation to achievment. *Studies in Second Language Acquisition, 6,* 54-68

Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly, 25,* 273-295.

Sarig, G. (1987). High-level reading in the first and in the foreign language: Some comparative process data. In J. Devine, P.L. Carrell, & D.E. Eskey (Eds.), *Research in reading in English as a second language* (pp. 107-123). Washington, DC: Teachers of English to Speakers of Other Languages.

Swann, J. (1992). *Girls, boys and language.* Oxford, UK: Blackwell.

Wildner-Bassett, M.E. (1992). *The interaction of learning styles and strategies: Diversity in the foreign language classroom.* Paper presented at the annual meeting of the American Council of Teachers of Foreign Language, Chicago.

Wolff, D. (1987). Some assumptions about second language text comprehension. *Studies in Second Language Acquisition, 9,* 307-326.

Zoubir-Shaw, S., & Oxford, R.L. (1995). Gender differences in language learning strategy use in university-level introductory French classes. In C.A. Klee (Ed.), *Faces in a crowd: The individual learner in multisection courses* (pp. 181-213). Boston: Heinle & Heinle.

Authors

DOLLY JESUSITA YOUNG, Associate Professor, Romance Languages, The University of Tennessee, 601 McClung Tower, Knoxville, TN 37996. (423) 974-2311. DJYOUNG@UTK.EDU. Specializations: foreign language education, second language acquisition.

1072

REBECCA OXFORD, Associate Dean, College of Education, The University of Alabama, 201 Carmichael Hall, Tuscaloosa, AL 35487. (205) 348-6051. BITNET: ROXFORD@BAMAED.UA.EDU. Specializations: foreign language education, learning styles and strategies.

1073

Applied Language Learning
1997, Vol. 8, No. 1, pp. 75-109

Pedagogical Intervention and Pragmatic Competence Development

Virginia LoCastro
International Christian University

The purpose of this exploratory case-study is to investigate the extent to which pedagogical intervention can facilitate the development of the pragmatic competence in English among foreign language learners. First-year university students enrolled in an intensive English-language program in Japan tape-recorded group discussions at the beginning and end of a nine-week term during which attempts were made to increase their level of awareness of and expand their ability to use markers of politeness in English. Utilizing a discourse/pragmatic analysis perspective, the data were analyzed for possible effects of the lessons. The results did not indicate the anticipated positive change in the learners' language behavior. Discussion focuses on the linguistic resources the learners did use to construct the speech event and on suggestions for the absence of linguistic politeness in the data examined.

Cross-cultural contact situations are frequently sites of negotiation, not only of propositional meaning of utterances but also of pragmatic meaning. Conventional as well as conversational implicatures give rise to the production of and reinforcement of stereotypes and possibly negative generalizations by the participants of each other's cultural identities.

One common concern in such interactional settings centers on politeness, or the lack thereof, defined as situationally appropriate behavior, a showing of attention to one's conversational partner's face needs.

In Japan, anecdotal evidence reports that Japanese hold the belief that "Westerners prefer directness." From the perspective of a speaker

1074

of Japanese, where linguistic politeness, or its absence, is obligatorily encoded, it is not unreasonable to mistake the fewer formal correlates of politeness in English-language discourse for a lack of the phenomenon. Both informal observation and the research literature (see, for example, Takahashi and Beebe, 1993) report that Japanese speakers are perceived by native speakers of the language as not being appropriately polite when using English.

Drawing on the work of Goffman (1963), who explored the disturbing effect on the social order when expected, "normal" reactions are violated or absent in interactions, a major assumption underlying this paper is that pragmatic competence development in language education should not be ignored. Linguistic politeness is only one constituent of pragmatic competence, yet a major one and inherently implicated in achieving successful communication in intercultural contact situations.

LoCastro (1994) has reported elsewhere on a study of textbooks used in senior high school English classes in the Japanese public school system, examined for evidence of attention to the teaching of politeness markers in junior and senior high schools. The Ministry of Education-approved textbooks provided few indications that Japanese learners of English are directly taught about politeness. Since reading and writing have been heavily emphasized over oral skills in Japanese English-language teaching (ELT), and, moreover, because few students have had role models of fluent speakers of English in their classes, it is not unreasonable that Japanese learners and speakers of English find it difficult to achieve their communicative needs of using situationally appropriate language.

The aim of the study reported on in this paper is to assess the results of an attempt to intervene in the acquisition of pragmatic competence by explicitly teaching about politeness [1] and conventionally polite forms during a nine-week term of an intensive language program. As it appears to be the case that linguistic etiquette is taught in first language acquisition contexts (Blum-Kulka, 1990; Snow et al., 1990), it seems reasonable to claim that the direct teaching of this aspect of pragmatic competence is needed (Ellis, 1992) in second or foreign language learning environments.

This is a study of naturally occurring talk; rather than explicitly eliciting language use which would include politeness markers, the task devised for the group discussions was assumed to lead to naturalistic use of such language behavior. While acknowledging that structured classroom group discussions are not an ideal sociocultural environment for collecting samples of language use which show a range of politeness

1075

features, this is a limitation of classroom-centered research in general and, despite the caution that must be used in generalizing from this study, useful insights may be achieved. Before proceeding to the study itself, it is necessary to (1) define politeness, (2) consider transfer of training, and (3) describe the analytic perspective to be taken.

Politeness is defined as the use of situationally appropriate language; specifically, it is viewed "as a strategy (or set of strategies) employed by a speaker to achieve a variety of goals, such as promoting or maintaining harmonious relations" (Thomas, 1995, pp. 157-158). One set of such strategies is the use of conventional politeness strategies (Thomas, 1995, p. 158); others include those proposed in Brown and Levinson (1987) and Leech (1983). Therefore, following Thomas (1995) and Green (1989), I adopt the view that politeness is to be understood as a "pragmatic/communicative phenomenon" (Thomas, 1995, p. 178) related to strategic, communicative goals of the participants. Researchers utilize various terms to label politeness phenomena, although none are without problems; nevertheless, these terms, in particular "linguistic etiquette," "linguistic politeness," and "politeness" shall be used as synonyms throughout this paper.

The present study explores what linguistic resources the participants in a particular speech event use to convey politeness. While in Japanese linguistic etiquette is obligatorily encoded in the language, in English the enactment of politeness is much less grammaticalized and prosodic features play a prominent role. Consequently, there are major structural and typological differences between Japanese and English with regard to linguistic politeness (see Niyekawa, 1991). This would presumably lead to difficulties in the enactment of politeness, precluding heavy reliance on transfer from the first language to the second language. Nevertheless, this study presupposes that the value attached to the need to communicate politeness would remain the same; irrespective of the language, of the sociocultural background of the participants, or of the speech event under observation, there is an underlying communicative need of human beings to achieve comity (Leech, 1983; Aston, 1993) in interpersonal contexts.

In order to address the learners' presumed need to use their second language effectively in communicative contexts, the present study tries to evaluate the effect of instruction by homing in on one type of transfer, namely, transfer of training between classes within an intensive English-language program. In the literature on transfer, Kasper locates this particular type of transfer under the category of "non-instructional factors" (1992, p. 219) and states:

1076

> instruction can be assumed to have a major role in shaping learners' perceptions of what is and is not transferable at the pragmalinguistic and sociopragmatic level, and hence have an impact on pragmatic transfer in learners' communicative behavior.

As these comments demonstrate, Kasper's emphasis is on transfer from learners' other languages, whereas I wish to examine the extent to which instruction can improve the learners' pragmatic competence —in effect, to enable them to improve their linguistic proficiency so that they can comprehend and produce situationally appropriate target-language behavior (Odlin, 1989, pp. 76-68). Rather than attempting to assess transfer from the classroom setting to non-instructional contexts, this study examines the informants' ability to utilize instructed linguistic behavior from a skills-training class in a skills-using class within an intensive language program.

Language educators regularly confront one of the major questions underlying much research in second language acquisition (SLA); i.e., can second or foreign languages be taught? Clearly, teachers of language hope to see positive transfer of training from the classroom to communicative situations in non-instruction environments. The literature on SLA (see Larsen-Freeman and Long, 199i and Ellis, 1994 for overviews of SLA research) documents the work which has been carried out specifically on this kind of transfer of training and how to increase the chances of it from occurring, particularly in predictable ways. Nevertheless, Ellis (1987, p. 179) asserts that the results of pedagogical intervention are "far from straightforward."

The type of transfer of training which is the current concern is limited to the carry-over from one type of class to another within the same intensive language program. It is not uncommon to find examples of negative transfer of training where learners exhibit forms in their interlanguage which derive from teachers having incorrectly understood and then taught a particular linguistic item. White (1993) cites the overuse of "please" taught as the translation equivalent of *doozo* in ELT classes in junior and senior high schools in Japan, resulting in mismatches of the illocutionary force of some speech acts and the linguistic form. Larsen-Freeman and Long (1991, p. 325) make the point that "it is reasonable to expect that formal instruction may trigger such processes as transfer, transfer of training and (over) generalization, depending on the choices teachers and materials writers make in this area." In the case of the present study, the focus is on positive transfer of the teaching of linguistic politeness items which constitute part of a learner's pragmatic competence of the target language across different classes

1077

within one language program. A language program must explore just how that can be done—that is, the kind of pedagogical intervention needs to be clarified and facilitated—and it is to that end that the present study was undertaken.

With regard to the analytical perspective, the data collected of classroom group discussions are analyzed for evidence of an increase in the ability to use designated speech acts and co-occurring redressive behavior by Japanese learners of English.

Specifically, this study looks at language use in three environments: (a) requesting answers from group members to questions on the task sheets; (b) directing the talk; and (c) seeking agreement/disagreement. [2] It was decided to limit the analysis to these three language functions on the rationale that they are more likely to engage the speakers in using face-threatening acts (FTAs). All three speech acts are potentially face-threatening to both speaker (S) and hearer (H). Given the definition of politeness suggested above, it is predictable that the speakers will attempt to use situationally appropriate (i.e., polite) linguistic forms, in order to achieve the perceived communicative goals of this speech event (i.e., carrying out a review and discussion in English of a reading assignment) using questions on task sheets as prompts. Acquiring the ability to use a greater variety of speech acts is not directly related to linguistic politeness. Nevertheless, the possibility of being able to choose and select from a variety of formal means to carry out linguistic action appropriate to the context of situation (e.g., suggest a contribution to the discussion rather than make a request) increases the learners' resources and awareness of and control over their own language behavior.

Method

In one type of lesson (Speaking or Sp) in the intensive language-training program, conventionally polite expressions and alternatives were explicitly taught. In those lessons, the intention was to raise awareness and to expand the learners' pragmatic competence. It was anticipated that there would be indications of an increase in awareness of politeness in a related type of lesson (Reading and Discussion skills or RD), in which an increase in the use of politeness markers in the interlanguage of the informants is considered to indicate an increase in awareness of the need for politeness, at least in the context of the speech event in question. While an increase in the use of politeness markers does not necessarily indicate greater awareness of linguistic politeness, it is assumed that there is more than a random possibility that such a correlation will occur.

1078

Following Brown and Levinson (1987, p. 22) and Schegloff (1993), I decided against statistical analysis as quantitative measures of politeness markers are intrinsically difficult for a number of reasons. According to Schegloff, quantitative analysis of talk-in-interaction fails to address such questions as what constitutes the occurrence of a linguistic behavior, how the counting takes into consideration the relevant linguistic and non-linguistic environments, and how quantitative analysis can characterize "the domain or universe from which our data are drawn" (1993, p. 110).

Further, the fact that the data of the present study consist of learner language (which raises the issue of defining consistently identifiable units in ungrammatical, disfluent talk) led to the conclusion that qualitative analysis would result in an admittedly different but, in my view, equally rich picture of the learner language being observed to that which a quantitative perspective would provide. Consequently, as my intention is to look only for a relative increase in awareness of politeness as a component of pragmatic competence, my analysis does not focus on statistically significant carry-over from one type of lesson to the other. Rather, the goal prioritizes insights to inform and motivate further studies of learners' pragmatic competence development.

The research question for this study is:

Is there transfer of training from one class to another of explicitly taught linguistic aspects of politeness?

The answer to the question entails an examination of the data for the speech act realizations as well as the means to maximize the efforts to achieve perceived communicative goals within the context of the speech event. As the program is an intensive, integrated one, one of the premises of the curriculum is that carry-over from one type of class to another is expected and to be facilitated. Therefore, it was judged that data collection in another class within the program, which specifically focuses on RD skills development, a likely locus of transfer from the Sp class lessons on communicative strategies, would be an appropriate site to learn about the effect of the treatment on the learners' pragmatic competence development.

In order to evaluate the effect of the pedagogical intervention, I examine the interlanguage performance of Japanese learners of English in the speech event of group discussions in an EFL classroom in Japan. Specifically, in the analysis of the data, focus was on speech acts and redressive linguistic action utilized by the informants. Drawing basic notions from the coding scheme in the Coding Manual of the CCSARP

(Cross-Cultural Speech Act Realization Project) project of Blum-Kulka, House, and Kasper (1989, pp. 273-294), my analysis focuses on the linguistic means, i.e. the surface forms of illocutionary acts, used by the learners to handle FTAs of requesting answers to questions on the task sheets, directing the discussion, and seeking agreement/disagreement. Note that, while the CCSARP coding scheme was developed for quantitative data analysis, the motivation for a modified use of the scheme in the present study derives from the need for a principled analysis of the audio-taped data without adopting the analytic perspective of the CCSARP. My rationale is that a study of how the informants carry out these three functions, examples of local phenomena, provides an indication of their ability to vary their language use more generally according to the weight of the imposition and the speaker-hearer relationship.

This is the unit of analysis:

HEAD ACT (+ alerters) (+ supportive moves)

The Head Act is the minimal unit; here is an example from Blum-Kulka, et al. (1989, p. 275):

John, get me a beer, please. I'm terribly thirsty.

The Head Act consists of the requesting speech act: get me a beer. Alerters, which tend to occur before the Head Act, include such elements as some formulaic expressions (e.g., "excuse me"), titles, names, address forms, etc. They are designed to get the attention of the addressee. Supportive moves, which can precede or follow the Head Act, are viewed as either external to the actual speech act, functioning as means to aggravate or mitigate the illocutionary force, such as giving a reason. There are also elements which act as internal modification devices, such as the use of conventional politeness markers like "please."

In addition, according to Blum-Kulka et al., a certain type of supportive move, hinting behavior, in the absence of a Head Act, can actually substitute for it. That is, as a nonconventional means to signal speaker intent, a hint can be regarded as functioning as a speech act of requesting. For example, in "The temperature has really dropped; it's getting awfully cold in here," the conversational implicature may be that of a request that a heater be turned on,[3] the specific illocutionary force assigned to it by the hearer being dependent on the speaker-hearer relationship as well as contextual information relevant to that particular occurrence.

1060

Although the coding scheme of Blum-Kulka et al. (1989) was designed for their cross-cultural study of the speech acts of requesting and apologizing, and they examined degrees of directness and indirectness as well as politeness features, I have adopted, with some modification, their basic framework for analysis of the interlanguage data of the group discussions. Instances of requesting answers of other group members will be examined for the kind of language used (i.e., declarative, imperative, or interrogative sentence forms) in the Head Act, and then any alerters and/or supportive moves are added to the minimal unit. An unmodified (i.e., no alerter or supportive move) speech act of requesting in the imperative form will be considered for the purposes of this study to be lacking as situationally appropriate language, following Brown and Levinson's (1987, p. 24) basic claim that "the kernel idea of our politeness theory (is) that some acts are intrinsically threatening to face and thus require 'softening.'" If the face threat of the request is not modified by any forms of redressive action, the hearer may withhold cooperation.

In sum, in order to operationalize the construct of situationally appropriate behavior, I regard alerters and supportive moves (including hints) as means to carry out redressive action. Clearly, although politeness does not necessarily correlate with linguistic form (see also Thomas, 1995, p. 157), it cannot be denied that the linguistic form of a speech act has a role to play as part of a complex of factors, summarized by the equation, linguistic form + context of utterance + speaker-hearer relationship, as suggested by Thomas (1995, p. 157).

Subjects

The informants in the study were all Japanese first-year university students at a liberal arts college in Tokyo. The average TOEFL score of entering freshmen is 490, with an average exit score after two years of 540. Within the context of the intensive English program, they participated in an intermediate RD skills class, which meets three times a week for 70 minutes each session. In addition, they had a communicative strategies or Sp class once a week, also 70 minutes long.

The two groups of informants (Groups A and B) had 21 students each (n=42). They were all between 18 and 23 (average 19) years of age when they participated in the project. Their responses to a background questionnaire they completed indicate that most had their previous training in the English language in Japan. Only five of them had spent one year or more outside of Japan for other than a short vacation or homestay experience. In two cases, they had spent one academic year abroad as exchange students in American high schools and three had lived abroad

1081

for 2.5 to 5 years. However, they had attended Japanese schools and lived with their own families, presumably therefore functioning in Japanese most of the time.

Two teachers were responsible for the RD classes, with two others for the Sp classes for the two groups. Group A started out with 21 and 18 remained at the end of the term. There were seven men and 14 women at the beginning of the term, with six men and 12 women at the end. Group B remained intact with the 21 students until the end of the term; six of them were men and 15 of them were women.[4]

Materials

The procedure to collect the data for this study consisted of two 30-minute audio-taping sessions, one at the beginning of the term (thus a "pre-test") to provide a baseline for the purpose of comparison, and a second taping at the end of the term (a "post-test") to check for evidence of transfer of training and an increase in awareness of linguistic politeness in English. During the nine-week term, the treatment (i.e., the explicit teaching of linguistic politeness) was carried out. Specifically, in the first week of the nine-week term, on the second day of classes and before the first speaking class had met, group discussions were audio-taped in the context of the RD classes. The students were asked to form groups of four or five and to answer and discuss questions on a prepared handout related to a homework reading assignment they had been given (see Appendix A). The topic was "educational values," something about which it was felt they should all have something to contribute, even if they had not carefully done the reading assignment. The questions were designed to review the content of their reading assignment and to provoke discussion on the topic in general.

The informants were asked to choose a group leader to elicit answers and to ascertain that every member of the group had a chance to participate. In addition, each group was asked to rotate the role of group leader among members. An argument was made that more politeness markers would be present in the speech of the group leaders, given the speech acts they would be using in the interaction. Due to the nature of the role of group leader, that person is likely to have to use such speech acts as requesting answers, directing the talk, and seeking agreement, all of which are potentially FTAs.

In addition, rotating the group leader is not uncommon as a pedagogical practice. Each group leader was asked to be responsible for soliciting answers to two or three of the 10 questions on the handout. The same data-collecting procedure was conducted during the last week

1082

of the term. The same task type was set. This time, however, the list of questions centered on two short stories which the students had read, thus again asking them to review the assignment and to discuss related topics (see Appendix B).

At both data-collecting times, each group included not more than one or two men and the informants were asked to avoid having "friends" in the same group. In both sessions, the students were asked if they minded being taped. They were told I was interested in learning to what extent they use the formulaic expressions they are taught in the speaking classes. Politeness markers were not mentioned explicitly. Due to the learner-centered, cooperative learning ethos of the language program, strong intervention in directing the make-up of groups was avoided. Note that tape-recording was judged to be the least intrusive and most likely methodology to provide natural speech samples of learner talk. The teachers did not participate in the discussions and were present in the classroom only to monitor the activity. The average length of taped discussions was 26 minutes, five seconds (total amount: 487.35 minutes). The recordings of nine of the 10 groups were used because one group did not record its discussion due to a technical failure during the first data collection session.

This was the apparatus of the study. I did not control for other possible variables, as I was interested in an aggregate or composite picture of progress in the use of linguistic markers of politeness in the learners' speech.

Procedure

During the nine-week term, the two Sp teachers explicitly addressed linguistic politeness in their lessons. This was done in several ways: by stressing the polite phrases in the list of useful expressions which the students received for each lesson (see Appendix C), by having the students highlight the relevant formal markers; by drawing their attention to intonation contours, facial expressions, gestures, and eye contact; by focusing on vocabulary choices; and by suggesting levels of politeness for addressing friends, teachers, etc.

The teachers commented on polite and impolite ways of expressing ideas, opinions, etc., and they emphasized these by saying: "A polite way to say/ask would be ..."

Examples of two particularly important lessons for this research project were those which focused on learning how to lead a discussion and give opinions and on language behavior when visiting a teacher's office. In Appendix D, the Sp lesson for the second week of the term,

1083

titled "Leading a discussion," is given. Below are one teacher's notes, directly quoted, for the lesson.

> Lesson Two: Leading a Discussion and Giving Opinions
>
> Emphasized being polite and letting everyone have a turn, trying to help those who can't express themselves (e.g., "maybe you mean," "if you don't mind, I think you mean...")
>
> Tried to show what would be rude: "that's wrong," "I don't understand you"—or ignoring the person when he/she speaks.
>
> Emphasized the need to try to get everyone to speak by asking "would anyone else like to ...?"
>
> Discussed agreeing — intonation, eye contact, and vocabulary—"Yes, you're right," "I agree."

Further comments from the Sp teachers' notes indicate they focused in particular on speech acts for asking for clarification, showing agreement/disagreement, eliciting another's opinion, and making suggestions. The learners practiced using "maybe" and eliciting an explanation for a statement of agreement.

In addition, the teachers worked on "asking for permission" or making "interruptions" as well as other speech acts as they came up in lessons.

Lastly, they would comment on cultural differences in language and other behaviors associated with politeness.

Results

A comparison of both sets of tape-recorded data indicates that the answer to the research question concerning transfer of training appears not to be clear. That is, evidence of transfer from the teaching of politeness in the speaking lessons to the group discussions in the reading-discussion classes was found, but not to the anticipated degree. The aim of this section then is to provide a comparison of the pre-test and post-test data to describe any changes in the use of surface level politeness phenomena.

1084

In the data labeled "pre-test," the Head Act for the requesting for answers frequently found was the inclusive imperative:

OK, let's start.
Let's discuss ...
Let's move to ...
Let's move on to ...
Let's go to the next one.

As for the most common Head Act in the context of directing the talk, there also tends to be one form, the interrogative, with little variation:

How about your opinion?
How about you?
What do you think (about that)*?*

When there was a noticeable silence signaling no group member was willing to talk, one person, typically the group leader, might say:

Does someone have an idea?
Do you have any ideas?

As for seeking agreement/disagreement, there were many more examples of ascertaining that the group members were in agreement than attempts to provoke disagreement. The speech act tended to consist of a Head Act without any alerters or supportive moves, as seen above with requesting and directing behaviors. [5]

No. 1

F2: *Do you agree with him?*
F3: *Yes, I hope so.*
M: *Me, too.*
F2: *Shall we go back to No. 5? We couldn't discuss enough the question.*

No. 2

F: *Anything that you think ... but the answer is what he said ... everyone thinks in your way, everyone agree with him, yes?*

1085

No. 3

M: *So learning how to think is intellectual way of thinking, right? So...*

Note that the linguistic form for seeking agreement shows some variation; in No. 1 interrogatives are used and in No. 2 and No. 3 there are summary statements followed by agreement-seeking tags. The relatively higher frequency of agreement is partially due to the pattern of the group leader's seeking consent from the other members that they all accept that the discussion can move on to the next question, a conversational management strategy not uncommon in talk in Japanese-language group discussions (Watanabe, 1993).

No. 4

M1: *Finished? I agree with you.*
M1: *Do you?*
M2: *Agree.*
M1: *Let's speak.*
F1: *Please go ahead.*
M1: *Go ahead.*
F2: *OK, may I read off?* (Reads next question on the task sheet)

As for disagreement, generally a dispreferred second-pair part of an adjacent pair (Pomerantz, 1984), there are predictably few tokens in the data. Here are two examples of showing disagreement with the previous speaker's talk:

No. 5

F1: *Do you agree with him?*
F2: *So do you mean you agree with him?*
M: *Partly I agree with you.*
F2: *I see.*
F1: *What point do you ...?* (unclear)
M: *By the way, please, some more saying about your opinion ...*

1086

No. 6

M: ... (deleted) ... *so it is good for him*
F2: *I don't think so.*
M: *What? Why?*
F2: ... (deleted) ...
M: *Objection.*
F2: *OK.*

In these two examples the disagreement is on-record (Brown and Levinson, 1987), with the addition in No. 5 of a form to mitigate the force (partly). However, there are some examples of utterances in which the likely conversational implicature is that the speaker disagrees in some way with what the previous speaker said.

No. 7

M: *I hear you don't have personal goals. Why did you take exam for ICU?*

No. 8

M: *You are leader.* (Three utterances deleted.)
Discussion is what effects should be changed to make college life more successful, so we must discuss about what should be changed.

The picture of the learner language at the start of the first term consequently suggests the following generalizations:

A. With regard to requesting behavior, the most commonly used form is "Let's" + verb. Thus, a minimal unit only—that is, a Head Act — is employed to realize a request for group participation in the speech event. There is speaker-hearer dominance—that is, both the speaker and hearer are included—by using "Let's ..." (Blum-Kulka et al., 1989, p. 278). There are no alerters or supportive moves; only the first person plural imperative form is used.

B. As for the linguistic means to direct the talk, again there is adherence to one means to elicit talk: an interrogative sentence form, asking for the participants' opinion or ideas about the propositional content of the previous speaker's talk. The interrogative tends to be a WH question ("what" or "how"), with a yes-no question type a less frequently used variation.

C. Seeking agreement in the two environments in which it occurred tends to be composed of a single Head Act as well, with no alerters or

1087

supportive moves. The syntactic forms showed more variation; declarative, interrogative, and imperative forms are found as realizations of seeking agreement. However, there are examples in the data I examined of what I will label a supportive move — specifically, a kind of grounder in which the speaker gives an explanation or summary of the previous talk, adding a tag question ("Yes?" "Right?") as a means to attend to the hearer's face needs, a bid for cooperation.

D. As for disagreement, this speech act predictably involves more language, given its nature of being a dispreferred act. In No. 5, M mitigates the FTA by choosing to say "I partly agree with you," using an adverbial modality marker and the positive form of the verb, rather than stating "I disagree with you." The second example also avoids the face threat such a statement would incur; "I don't think so" places full responsibility on the speaker, thereby lessening the possible negative impact. This interpretation could also be said of the speaker's use of the word "objection." Although it is lacking in any form of linguistic mitigation of the speech act, the fact that the hearer is not mentioned makes this a speaker-dominant utterance, thereby implicating the speaker and not the hearer. In this example, the pragmatic meaning or conversational implicature makes the speech acts of disagreement less impolite than analysis of the linguistic form alone would give.

In general, all of the linguistic realizations of the speech acts in question tend to be composed of the Head Act alone, without alerters or supportive moves—i.e., without the linguistic forms which serve to carry out redressive action of the FTAs.

It is necessary to compare this picture of the learners' use of language in these three contexts with that in the data in the post-test. Again, using the sequence of Head Act (+ alerters) (+ supportive moves), the analysis focuses on evidence that the explicit awareness-raising and teaching of politeness had an effect and then at the characteristics of the observed effects. Some changes can be observed.

At the time of the post-test, in the first category of requesting behavior, the basic unit, the Head Act, is modified by more language in the form of alerters and supportive moves. There are numerous examples of some form of address term: surname, first name, or nickname, sometimes with a Japanese suffix, either + *san* or (rarely) + *chan*.

> *Aki-san, what would you use to describe him?*
> *How about you, Okura-san?*
> *I'm sorry I can't discuss. Kyoko?*

1088

The other observed change with requests is that the solicits are longer, incorporating lexis used in the task sheet. Rather than the more general requests ("What do you think?" or "Let's start.") found in the pre-test data, now the request realizations are more focused, thereby providing a non-conventional means of, in my view, more situationally appropriate language.

> *And what support do you did you find in this story?*
> *Do you have another possible scene?*
> *What adjectives would you use to describe him?*

These utterances cannot be said to include alerters or supportive moves. Nevertheless, they do attend to the hearer's face needs by attempting to be more explicit in communicating the propositional content.

As for the language used to direct the discussion, the following is noticeable:

A. There is more frequent use of modal auxiliaries with interrogatives, all conventionally signaling redressive action:

> *Can you answer this question?*
> *What adjectives would you use to describe him?*
> *Can you provide evidence?*
> *Shall we go on?*

B. The solicits may be longer and more specific:

> *So what do you think about his last action?*
> *So how how can it be suspense for you?*

> *So No.6* (reads question from the task sheet),
> *it's a interesting point, I think.*

C. Greater use of supportive moves can be found:

> *Yeah, it's same, but that's a point, a good point, how about you, Midori?*

> *Let's go to next question* (reads from the task sheet). *Are you ready to answer this question, Eiko? How about Mutsumi?*

1089

In the first example here, the speaker gives a reason for agreeing. In the second, the hearer is given the option to refuse a turn at Sp. Both can be labeled supportive moves.

Similar modifications in seeking agreement/disagreement are also present in the data examined.

A. The language used to seek agreement, i.e., to check that all the group members are ready to move on to the next question, can be observed in the following example:

No. 9

F: *I agree with her.*
M: *Yes, I agree.*
S: *Yes, everyone think in that way.*
M: *That's a kind of common opinion.*
S: *Yes, that's we have a usual opinion.*
F: *What is the ...* (unclear)*?*
S: *Say yes. OK, let's move on.*

This is an excellent example of the joint creation of talk, where the three participants contribute until S establishes at the end of this sequence that they can move on to discussion of the next question. Had S stated that wish at the beginning before F and M had had time to talk, S's behavior would have been viewed as less than polite by her conversational partners. It is possible to claim that this joint creation of talk can be labeled situationally appropriate talk, with each turn at talk contributing a supportive move to the Head Act in the final turn of S in this sequence.

B. There appear to be more frequent challenges of the previous speaker's talk and more supportive moves are used. These speech acts of disagreement take the following forms:

No. 10

Naoko: *He confident he has confidence on his ability.*
Kumi: *But which part shows that same opinion that he* ... (unclear).

1090

No. 11

H: *OK, what is your opinion?*
N: *I, I agree with he—um—I can imagine it easily*
H: *Really? Why, why can you imagine it easily?*
N: *Because ... um.*

No. 12

M: *So the need for oral and written stories is decreased, uh, decreasing.*
A: *Why do you agree?*

The second speakers' utterances in these examples, given on record with no mitigation, suggest attempts to apply one of the principles of critical thinking, a salient aspect of the curriculum of the English-language program. The informants sought to push the discussion beyond simply having each person present his/her own comments without any development of the topic, consequently producing situationally appropriate language behavior

In summary, several changes in the learners' interlanguage with regard to the use of situationally appropriate language can be found:

1. The linguistic forms of the Head Acts show more variation.
2. Alerters are observed, specifically forms of address.
3. Internal modification by means of modal auxiliaries is present.
4. More details are added to make the propositional content, as well as the illocutionary force, more focused, i.e., more supportive language behavior, although interpretation of speaker meaning is dependent on inferring from essentially non-conventional means to convey consideration of the co-participants' face wants.

Discussion

In the previous section, the analysis of the data provided a view of the extent to which the informants used linguistic markers of politeness in their enactment of the speech event in three specified environments. It has to be acknowledged that the data I examined did not provide evidence that the pedagogical intervention had a noticeable effect on the learners' ability to use linguistic etiquette. Consequently, the research question concerning evidence of transfer of training from one class to

1091

another has to be answered negatively. As the studies reported on in Lightbown and Spada (1993, pp. 90-96) indicate, efforts to intervene in instructed SLA can go wrong for any number of reasons, including poor teaching, insufficient time for the effect to become noticeable within the context of an empirical study, and the influence of developmental stages according to the learnability hypothesis (see Pienemann, 1984), which could lead to a claim that there may be constraints on the learning of such features of language as those which correlate with linguistic etiquette (see Kasper, 1979). In this section, some reasons for the lack of evidence of transfer are discussed, followed by comments on aspects of the sociocultural context, often ignored in similar studies, as it is my view that it is unwise to avoid attention to influences from the context in which the learning setting is embedded. Speech acts are highly sensitive to situational variables (Blum-Kulka et al., 1989).

First of all, the data may reflect the learning which derives from materials and instruction in junior high school English classes where students typically memorize formulaic expressions as unexamined chunks of language, limiting the possible range and creativity of their expressions. In senior high school, most of the time is spent on reading and translation skills development, with little or no attention to pragmatic competence development. Japanese informants claim that they only learn "thank you" and "please," to be used in any and all situations where there is a need to signal politeness. The formulaic chunks of language tend to be used as politeness markers in any or all contexts, an example of overgeneralization as a communication strategy.

Consequently, given the lack of attention to pragmatic competence development in their previous language-learning experiences and, in addition, the fostering of dependence on formulaic expressions, slotted-in as needed without regard to the pragmatic equivalent of the phrase or item, it is not unreasonable to find the first-year university learners in this study resorting to overuse of the expressions on class handouts at the beginning of the first term. The transcription of the pre-test data does indicate dependence on class materials (see Appendix F for one example from a Sp lesson).

Nevertheless, as the analysis in the Results section shows, some traces of transfer of training from actual instruction can be seen. In particular, there are several instances where the students seem to have understood the need to mitigate their utterances, yet the politeness markers had not been learned with accuracy. That is, they had not achieved mastery or full control over use of the form.

1092

> M: *I was also surprised ... same opinion.*
> E: *I also think, I, I also don't agree ...* (deleted) ... *so I don't agree that.*

These two contributions, which are second-pair parts of adjacency pairs, provide evidence in the repetition, especially of "also," of some effort to go beyond a formulaic, agreement speech act response, yet both attempts are abandoned. The teaching in the Sp classes may not have resulted in learning to the level of accuracy and automaticity. Retrieving knowledge and making it accessible for use in communicative contexts is problematic for learners in EFL contexts (Bialystok, 1993). The one clear indication of learning is the increase in the use of using one's conversational partner's name—direct address—before asking for an opinion, for example. Including the interlocutor's name can be a politeness marker and there is evidence that the learners had noticed use of that behavior in the classroom context by the teachers and were able to use it.

It may be that a nine-week term is not long enough for transfer of training to be observed to a significant level. A longitudinal study of some duration might generate sufficient evidence to enable researchers and teachers to begin to predict when intervention could occur to facilitate learning of certain aspects of pragmatic competence. Note, furthermore, that the students knew that they would not be tested on what they had learned in the Sp lessons; needless to say, knowledge of being tested on the language content of the Sp lessons might have resulted in more transfer of training to the RD classes.

Aspects of the particular sociocultural context are addressed which need to be considered in order to develop a more inclusive model of the variables influencing the foreign language learners' pragmatic competence. In my view, formal or instructed SLA settings cannot be studied without regard to the broader societal environment in which they are embedded. I would suggest this is particularly the case for pragmatic competence assessment as it inherently engages the participants in interactive, interpersonal events, which are grounded in the social enactment of meaning.

With regard to sociocultural dimensions, one consideration is the extent to which the data reflect the informants' attempts to achieve comity (Aston, 1993) within the context of this speech event. A salient social variable of the group discussions is the fact that the learners are first year university students who would all be in the same group for the entire academic year for their English classes. As peers, that is, of about the same age and in the same year at their university, they might

1093

not have felt the need to use politeness. Wolfson (1986) claims that an established, clear social contact situation leads to less negotiation and presumably less of a need to engage in complex linguistic etiquette behavior. In addition, Ellis (1992) suggests that, as a result of his study of requesting behavior of two young learners in a classroom environment, basic, but not elaborated forms of requesting behavior are likely due to the routineness of the classroom tasks and interactions, and to the lack of a strong communicative need to motivate more complex requests. This interpretation appears highly plausible. However, no baseline data is available of what native speakers would do in similar classroom situations.

Yet another interpretation concerns the speech act of requesting. Although for the purpose of data analysis, three different language functions were distinguished—that is, requesting opinions, directing the talk, and seeking agreement/disagreement. It seems possible to argue that all three are subcategories of requesting behavior. Tsui (1994, see pp. 91-95) reviews the literature on such speech acts and claims that a salient feature to distinguish a request from an order or a command is the possibility that the addressee can opt out of carrying out the designated behavior. In addition, according to Tsui, the lack of optionality as well as the presence of a context in which cooperation is assumed from the co-participants of an interaction results in "I want you to do X" type statements ("wants and needs" statements, in Tsui terminology) being used without mitigation. As the data I examined were collected in a classroom environment where learners typically are expected to do as the teacher wishes, or the teacher's surrogate, in the form of the discussion group leader, redressive action would be less likely to appear in such a speech act environment and conventionally direct requests would be the unmarked form. Consequently, if all the speech acts analyzed above are viewed as subcategories of requesting behavior, and if the informant group leader presupposed cooperation from the others, there being little social distance between them (such as in families), it is reasonable that there would be few linguistic politeness markers. In their study of indirectness, a feature of talk which tends to co-occur with politeness phenomena, Blum-Kulka and House (1989) found variation in requests according to such features as likelihood of compliance, degree of difficulty in complying, and perceived social distance. Thus, the relative lack of linguistic etiquette in the data from the group discussions may be explained on the same grounds, reflecting the findings of other studies.

Finally, despite the fact that at the time of the first taping the informants had been in class together for only one day, thus making it

difficult to argue that they were "friends," recent literature as well as anecdotal evidence indicates that directness in such a setting is more tolerated by Japanese than by American or Canadian speakers of English (see Rose, 1992, among others). University students tend to be considered a separate sub-culture in many areas of the world and it is possible that the language behavior of the informants reflects practices of their sub-culture in Japan and elsewhere. Politeness enactment indicates a sensitivity to hierarchy between and among conversational participants (Hinkel, 1994). Consequently, the informants may have also been unwilling to use politeness markers frequently with others who are their direct peers, as a pro-active means to build solidarity among the group members. There is also the influence of the *uchi* (in-group) vs. *soto* (out-group) dichotomy, which is viewed as playing a strong role in interpreting Japanese interpersonal behavior (Lebra, 1976; Bachnik, 1994). Cross-linguistic influences may be operating and, despite the fact that the informants had only relatively recently become in-group members of the English class and of the university community as a whole, it is possible that they are perceived by each other as being in-group members and therefore the use of informal language between and among members of the groups commonly found in Japanese-language use (Niyekawa, 1991) is being transferred into the English language context.

Conclusions

This paper has reported on a study of positive transfer of training of linguistic markers of politeness. The pedagogical intervention seemed to have had little effect on the learners' ability to produce situationally appropriate language behavior more frequently in the context of the speech event of group discussion. It had been hoped that the direct teaching about politeness and of linguistic means to enact politeness would contribute to resolving the problem stated at the beginning of this paper, that is, that Japanese learners and speakers of English tend to be perceived as less than appropriately polite by native speakers of English. Therefore, the original problem remains.

From the point of view of theoretical implications, it appears that pragmatic competence, specifically linguistic etiquette, remains a fuzzy category not yet clearly enough understood to be taken on board by classroom practitioners. This is not surprising as the notion of politeness itself continues to be problematic even in discussions of its enactment in one language, English or Japanese. Communicative language teaching has been popular for over a decade, although only officially

1095

acknowledged by the Japanese Ministry of Education more recently; the term "pragmatic competence" has appeared in the discussions about what classroom practitioners are to do to facilitate development of communicative ability. However, because situationally appropriate language use taps into unexamined, unconscious values and beliefs, it is difficult to contemplate how teachers and learners are to navigate between the two poles, one being focus on form for accuracy and the other using language in communicative, information-gap type tasks with little or no teacher input on which conversational gambit to use when and with whom. If models of SLA are to be of use, they must go beyond the continued focus on sentence-level of grammar, specifically on phonological and grammatical aspects, to account for pragmatic meaning and for creative language use, producing cohesive, appropriate texts in a second/foreign language.

Some conclusions, nevertheless, can be drawn from the study so far.

First, pragmatic competence development is a complex phenomenon involving values and beliefs, level of language proficiency, and social norms and practices, all of which interact with and constrain each other.

Second, learners and speakers of a second/foreign language are not free of the influences of their previous language learning environments, which reflect societal attitudes towards language, their mother tongue, and foreign languages.

Third, SL development, particularly of psycholinguistically problematic aspects, requires more time than anticipated as well as more extensive exposure to and experience with naturalistic target language input (see Bouton, 1994).

The task for language educators is not a small one; yet the ability to achieve comity, and avoid conflict, in cross-cultural contact situations may depend on it.

Further Research

Clearly, further research is called for in order to address the question of whether pragmatic competence can be taught. The focus should not be on any one language community as in this case study. Rather, researchers should broaden the scope to include pragmatic competence development—that is, the ability to use language in situationally appropriate ways, in the first language as well as in second languages of other groups, and in both second- as well as foreign-language environments. Projects could require tape-recording of classroom lessons on

1096

politeness, collecting input from the teacher as well as output of the learners. In addition, retrospective interviews and informant diaries, as well as other accounts of learners asked to introspect about their communicative strategies and to notice contextual features of the instances when they use linguistic etiquette, would provide invaluable information.

It should also be noted that studies of this sort need to be done at the secondary school level, in the form of collaborative research with teachers, learners, and applied linguists working together from the start of language education in school systems wherever second/foreign languages are taught. Further, there is a need for comparison, baseline data of learners interacting in their mother tongues and in second/foreign languages in the same speech event and context.

Both quantitative and qualitative studies are needed. However, until researchers have come closer to solving the questions mentioned above, posed by Schegloff (1993), microanalysis of single speech events of talk-in-interaction are needed to inform language educators, teacher trainers, and materials and curriculum designers without delay.[6]

Notes

This is a revision of a paper given at the Second Language Research Forum 1994 in Montreal, Quebec, Canada. The author wishes to thank Tandy Bailey, Brenda Bushell, and Ged O'Connell for their invaluable help in collecting the data, transcribing the discussions, and in teaching the informants in the Sp and RD classes. This study would not have been possible without their cooperation and contributions. I would also like to thank three anonymous reviewers for their helpful comments and suggestions. It goes without saying that the remaining shortcomings are no one's responsibility but my own.

1. It is to be acknowledged that there may be no universal system of situationally appropriate language use, commonly termed "politeness," even for native speakers of English. Although this study does not focus on a particular native speaker (NS) system, as the author was raised in the United States, it can be claimed that the Japanese informants' language use is being compared to a generalized American system of positive politeness, following Scollon and Scollon (1983).

2. The learners are instructed in the English program to seek and express disagreement or to challenge rather than accept the ideas or opinions of other participants as one element in the critical thinking goals of the curriculum. Also note that with regard to speech acts of agreement and disagreement, in this case I am interested in the linguistic

1097

efforts made to solicit or seek agreement or disagreement in the first-pair part of an adjacency pair, not in any response in the second-pair part.

3. In the Discussion section, I comment further on definitions of requesting behavior.

4. The possibility of gender differences in the use of formal markers of politeness should ideally be considered. However, due to the small sample size (seven men and 14 women) it was decided a composite of language behavior of all the informants was appropriate.

5. Transcription Note:

... = hesitation/silence
( ) = inaudible or unclear
F = female participant
M = male participant
F1, F2, etc. = female informant 1, female informant 2

Note that in some of the examples, first name initials or actual names are used, although they have been changed to protect anonymity. The deleted utterances were judged unnecessary for the current study. I use the definition of utterance of Crookes and Rulon (1985, p.9): "an utterance [is] defined as a stream of speech with at least one of the following characteristics: (1) under one intonation contour, (2) bounded by pauses, and (3) constituting one semantic unit."

6. LoCastro and Netsu are currently engaged in a research project to secure relevant baseline data.

References

Aston, G. (1993). Notes on the interlanguage of comity. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 224-50). New York: Oxford University Press.

Bachnik, J. M. (1994). Introduction: *uchi/soto:* Challenging our conceptualizations of self, social order, and language. In J. M. Bachnik & C. J. Quinn (Eds.), *Situated meaning* (pp. 3-37). Princeton: Princeton University Press.

Beebe, L. M. & Takahashi, T. (1989). Do you have a bag?: Social status and patterned variation in second language acquisition. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.). *Variation in second-language acquisition* (Vol. 1: *Discourse and pragmatics*, pp. 103-128). Clevedon: Multilingual Matters.

1098

Bialystok, E. (1993). Symbolic representation and attentional control in pragmatic competence. In G.Kasper & S.Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 43-57). New York: Oxford University Press.

Blum-Kulka, S.(1990). Don't touch the lettuce with your fingers: Parental politeness in family discourse. *Journal of Pragmatics, 14*, 259-288.

Blum-Kulka, S. and House, J. (1989). Cross-cultural and situational variation in requestive behavior. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies*. Norwood, NJ: Ablex.

Blum-Kulka, S. House, J. & Kasper, G. (1989). *Cross-cultural pragmatics: Requests and apologies*. Norwood, NJ: Ablex.

Bouton, L. (1994). Conversational implicature in the second language: Learned slowly when not deliberately taught. *Journal of Pragmatics, 22*, 157-167.

Brown, P. and Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Crookes, G. V. and Rulon, K. (1985). Incorporation of corrective feedback in native speaker/non-native speaker conversation. Technical Report No. 3. Center for Second Language Classroom Research, Social Science Research Institute, University of Hawaii, Honolulu. Cited in Crookes, G. V. (1990). The utterance and other basic units for second language discourse analysis. *Applied Linguistics, 11*(2), 183-199.

Ellis, R. (1987). Contextual variability in second language acquisition and the relevancy of language teaching. In R. Ellis (Ed.), *Second language acquisition in context* (pp. 179-194). New York: Prentice Hall.

Ellis, R. (1992). Learning to communicate in the classroom: A study of two language learners' requests. *Studies in Second Language Acquisition, 14*,1-23.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gatherings*. New York: The Free Press.

Green, G. M. (1989). *Pragmatics and natural language understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hinkel, E. (1994). Pragmatics of interaction: Expressing thanks in a second language. *Applied Language Learning, 5*(1), 73-91.

Kasper, G. (1979). Communication strategies: Modality reduction.

1099

Interlanguage Studies Bulletin, 4, 266-281.

Kasper, G. (1992). Pragmatic transfer. *Second Language Research, 8*(3), 203-231.

Larsen-Freeman, D. & Long, M. H. (1991). *An introduction to second language acquisition research.* London: Longman.

Lebra, T. S. (1976). *Japanese patterns of behavior.* Honolulu: University of Hawaii Press.

Leech, G. N. (1983). *Principles of pragmatics.* London: Longman.

Lightbown, P. M. and Spada, N. (1993). *How languages are learned.* Oxford: Oxford University Press.

LoCastro, V. (1994). Linguistic politeness markers in the speech of Japanese speakers of English. Paper given at the Eighth Annual International Conference on Pragmatics and Language Learning. University of Illinois at Urbana-Champaign, March 31-April 2.

Niyekawa, A. M. (1991). *Minimum essential politeness.* Tokyo: Kodansha International.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning.* Cambridge: Cambridge University Press.

Pieneman, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition, 6,* 186-214.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. Maxwell Atkinson and J. Heritage (Eds.), *Structures in social action* (pp. 57-101). Cambridge: Cambridge University Press.

Rose, K. (1992). Method and scope in cross cultural speech act research: A contrastive study of requests in Japanese and English. Unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction, 26*(1), 99-128.

Scollon, R. and Scollon, S. B. K. (1983). Face in interethnic communication. In J. C. Richards and R. W. Schmidt (Eds.), *Language and communication* (pp. 156-190). London: Longman.

Snow, C. E., Perlmann, R. Y., Gleason, J. B. and Hooshyar, H. (1990). Developmental perspectives on politeness: Sources of children's knowledge. *Journal of Pragmatics, 14,* 289-305.

Takahashi, T. & Beebe, L. M. (1993). Cross-linguistic influence in the speech act of correction. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 138-158). New York: Oxford University Press.

Thomas, J. (1995). *Meaning in interaction: An introduction to prag-*

110

matics. London: Longman.
Tsui, A. B. M. (1994). *English conversation*. Oxford: Oxford University Press.
Watanabe, S. (1993). Cultural differences in framing: American and Japanese group discussions. In D. Tannen (Ed.), *Framing in discourse* (pp. 176-209). New York: Oxford University Press.
White, R. (1993). Saying *please*: Pragmalinguistic failure in English interaction. *English Language Teaching Journal, 47/3,* 193-202.
Wolfson, N. (1986). Research methodology and the question of validity. *TESOL Quarterly, 20,* 689-99.

Appendices

Appendix A

Pre-test Task Worksheet

College Thinking*

Introduction, Discussion, Questions

In small groups, discuss the following questions.

1. In the author's opinion, what aspects of college orientation should be changed to make college life more successful?

2. Discuss the orientation you have had at ICU. What aspects have been most useful and least useful for your future at ICU?

3. Why, according the author, are professors at many colleges disappointed by the standards of student writing?

4a. What does the author mean by "developing a framework in which to place the courses that you take in college?

4b. Why is it important to understand the framework of a college program? Do you think this is also true of a Japanese university?

5. Do you think it is important to establish personal goals for yourself as you begin your university career? Give reasons for your answer.

1101

6a. Professors emphasize that although learning the content is important, learning intellectual skills and attitudes is far more important. What does this mean?

6b. Is this philosophy different from what high school teachers tell you?

7. "Learning how to think in college will help you later in life." What does the author mean by this?

8. What is meant by "critical thinking"?

9. The author encourages students to challenge the instructor/professor and classmates. What must we do in order to feel confident in challenging our teachers and peers? (What steps must we take?)

10. How can we challenge each other and still show respect for one another?

---

* Meiland, J.W. (1981). College thinking: How to get the best out of college. New York: Mentor.

Appendix B

Post-test Task Worksheet

Under the Banyan Tree**

Discussion Questions

1. Nambi is the main character (who are the others?) of this story. Who is he? What adjectives would you use to describe him?

2. Why did he do what he did at the end of the story? Provide evidence from the story to support your point of view.

3. Compare Vera and Nambi. Both of them tell stories, but there are differences. What are they?

4. What happens in the story? Is there a building up of suspense? Were you surprised by the ending? Why/why not?

5. Describe the village. Can you imagine it easily? Why/why not?

1102

6. What are the themes in this story? For example, is "growing old" a possible theme?

7. If you look at this story from a cross cultural point of view, what aspects of the story do you notice?

8. "The cinema, and more recently television, have completely replaced the need of oral and written stories." Do you agree?

9. Is there a tradition of oral story telling in Japan? What is it? Are there similarities or differences between the tradition in Japan and what you can see in the Nambi story?

10. What was your reaction to the story? Give reasons for your opinion.

** "Under the Banyan Tree," by R.K. Narayan and "The Open Window," by "Saki" (H.H. Munro). (1992). In Alice Martin and Robert Hill (Eds.), *Modern short stories*. Hemel-Hempstead, UK: Prentice-Hall International.

Appendix C

List of Useful Expressions from a Speaking Lesson

Focus of Lesson

Visiting a Teacher's Office
Active Listening Skills

Useful Expressions

Opening a conversation

Excuse me, but do you have a minute?
I'm sorry to bother you, but do you have a minute?
Excuse me for interrupting you, but may I ask you a few quick questions?

Stating your business

I'd like to talk with you about...
I have a small problem I'd like to discuss with you.
I'd like to ask your advice on something.

1103

Closing a conversation

Well, thank you very much for your time.
Well, I know you're busy. Thanks for your time.

Active listening skills

Nodding the head.
Saying, *Uh huh. Oh really? I see.*
Making eye contact.
Taking notes.

Appendix D

Speaking Lesson

Leading a Discussion

Focus of Lesson

How to Be a Small Group Discussion Leader

Useful Expressions

1. Opening a Discussion
Today we need to decide/ discuss/prepare...
Let's begin with...

2. Eliciting an Opinion
What about...?
What do you think [name]*?*
Does anybody have anything else to add?
Would anybody like to add to what [name] *has just said?*

3. Responding & Summerizing
Yes, that's a good point...
Okay, so far we've said...
Okay, to sum up, we've said...

4. Closing a Discussion
Okay, so we all agree?/Do we all agree?
Let's move on to the next point.
Well, that's all we have time for.

5. Giving an Opinion
I think...
It seems to me...
As far as I'm concerned...

6. Agreeing
I agree.
I think so too.

7. Disagreeing
I see your point, but...

1104

Activity Two Worksheet (Student A)

Role One: Group Leader	Useful Expressions
You are the group leader. The topic for for discussion is whether or not women should work outside the home. Your responsibility is to keep the conversation going and to make sure that the other group members speak for an equal amount of time. Try to bring the discussion to a conclusion. Be sure to use the expressions on the right hand side of this page.	Opening a Discussion *Today we need to decide/* discuss/prepare... *Let's begin with...* Eliciting an Opinion *What about...?* *What do you think* [name]*?* *Does anybody have anything else to add?* *Would anybody like to add to what* [name] *has just said?* Responding & Summarizing *Yes, that's a good point.* *Okay, so far we've said...* *Okay, to sum up, we've said...* Closing a Discussion *Okay, so we all agree?/Do we all agree?* *Let's move on the next point.*

Role Two: Participant	Useful Expressions
You believe that all university students should be required to wear uniforms because it increases morale provides a sense of school identity helps students focus on their schoolwork rather than on their appearances is less distracting for teachers keeps costs down Be sure to use the expressions on the right hand side of this page.	Giving an Opinion *I think...* *It seems to me...* *As far as I'm concerned...* Agreeing *I agree.* Disagreeing *I see your point,but...*

Role Three: Participant	Useful Expressions
You believe that cigarette smoking should not be banned in public places. It should be a matter of choice because cigarette smoking allows for human individuality other things which are more harmful are not banned cigarette smoking reduces stress cigarette sales increase tax revenue	Giving an Opinion *I think...* *It seems to me...* *As far as I'm concerned...* Agreeing *I agree.* Disagreeing I *see your point, but...*

1106

Appendix E

RD Class Handout On How to Do a Discussion

How to Do a Discussion

First of all, choose a discussion leader and, if asked to do so, a secretary, who will take notes and then use those notes to give an oral report on your group's discussion to the whole class.

You can choose these two people in various ways. But it is important to remember that the same people should not always be the leader or the secretary.

The role of the leader is to (1) make sure everyone in the group has a chance to talk and express his/her opinions and (2) be non-judgmental, that is, not express his/her opinions and control the group. The leader needs to invite others to talk and not be too quick to say what he/she thinks.

Some examples of the *language* you can use.

Anyone:

Okay, who wants to be the leader today? And the secretary?
How about you, Hiroshi? Let's junken.
All right. Keiko is the leader. (Toshinori is the secretary.)

Keiko:

Let's get started. We're going to discuss... (state the topic/read the question, etc.).
Shall we just go around the circle and each person can say what he or she thinks?

Or

Mariko, would you like to start?
(Then Keiko has to ask each person to talk.)

If there is some disagreement/if the leader wants to get people to discuss different points of view:

Keiko:

Mariko thinks XXX. Kaoru, do you agree with her?

If there are several questions to discuss:

All right. Let's move on to the next question/the next topic.

Just before time is up, then the group needs to check with the secretary to see that the main points have been listed. You need to help the secretary be ready to report to the class.

Author

VIRGINIA LOCASTRO, Associate Professor, Division of Languages, International Christian University, 3-10-2 Osawa, Mitaka, Tokyo 181, Japan. e-mail: locastro@icu.ac.jp. Specializations: second-language acquisition, discourse analysis, pragmatics, sociolinguistics.

1108

RESEARCH REPORT

Applied Language Learning
1997, Vol. 8, No. 1, pp. 111-118

Free Voluntary Reading as a Predictor of TOEFL Scores

Rebecca Constantino
U.S. Congress

Sy-Ying Lee
National Taiwan Ocean University

Kyung-Sook Cho
Pusan National University of Education

Stephen Krashen
University of Southern California

Forty-three international university students, currently living in the United States, filled out a questionnaire probing years of English study, length of residence (LOR) in the US, free reading habits in the first and second language, and TV watching. Despite the fact that subjects reported little reading in English, this variable was a significant predictor of TOEFL test performance. In addition, English study in the home country and length of residence in the US were also related to TOEFL scores.

Hanania and Gradman (1991) reported that the best predictor of TOEFL scores among international students who took the exam before coming to the United States was "extracurricular reading." Constantino (1995), in a case study of four international students in the United States, also reported that increased pleasure reading had a very positive effect on TOEFL performance among the three who did extensive pleasure reading; her subject who did only formal study did not do well on the exam. These results are in agreement with a large number of studies showing a clear relationship between pleasure reading and several aspects of literacy development, including reading comprehension.

1109

© 1997, Rebecca Constantino, Sy-Ying Lee, Kyung-Sook Cho, and Stephen Krashen

writing style, spelling, grammatical competence, and vocabulary (reviewed by Krashen, 1993). In this study, we investigate the relationship between pleasure reading and TOEFL scores for international students currently living the United States.

Procedure

Subjects: 43 international students (17 female, 26 male) from 14 different countries (South Korea, China, Taiwan, Saudi Arabia, India, Switzerland, Spain, Japan, Brazil, Indonesia, Thailand, Lebanon, Ukraine, and Hong Kong) participated in the study. All subjects were currently university students in the United States, and some are still enrolled in English as a Second Language classes. All had taken the TOEFL examination.

Instrument: Subjects filled out a 13-item questionnaire, which included questions about their TOEFL score, the frequency and amount of reading done in their first language and in English, TV viewing in English, frequency of visits to the library and bookstore, and the number of books they read in English before taking the TOEFL examination. Subjects required a maximum of ten minutes to fill out the questionnaire.

Results

Table 1 presents the results of the questionnaire.

Table 2 indicates that TOEFL scores were correlated with amount of English study done in the home country ($r = .367$), the composite measure of free reading ($r = .39$), and the number of books read ($r = .447$). Additional study done in the US was not a significant predictor of TOEFL scores, but the mean amount of study done in the US was only 1.35 years. The correlation between length of residence in the US and TOEFL scores was positive but did not reach statistical significance.

Table 3 presents the results of a simultaneous multiple regression analysis. Only three variables were entered into the analysis, because of the relatively small number of subjects used. The variables used were those that were of the most theoretical interest, and that showed modest to high correlations with TOEFL scores. "Books read" and "free reading" were combined into one predictor, as they were highly correlated ($r = .714$).

1110

TABLE 1
Descriptive Statistics

Item	Mean	s.d.
1. Age	25.79	4.03
2. Score on TOEFL examination	564.55	49.35
3. Years of English study in home country	8.60	3.54
4. Years of English study in US before taking TOEFL	1.35	3.39
5. Length of residence in US (months)	18.74	22.06
6. Amount of English TV watched each day before taking TOEFL 1 = never 2 = less than one hour/day 3 = 1-3 hours/day 4 = more than 3 hours/day	2.09	.781
7. Frequency of English newspaper reading before taking TOEFL 1 = never 2 = less than one hour/day 3 = 1-3 hours/day 4 = more than 3 hours/day	1.70	.599
8. Amount of free reading done in English before taking TOEFL 1 = none 2 = 1-3 months 3 = 4-6 months 4 = 7-12 months 5 = 1-2 years 6 = more than 2 years	2.84	1.70

1111

TABLE 1 (cont.)

Item	Mean	s.d.
9. Frequency of reading books for fun in English before taking TOEFL 1 = none 2 = 1-2 times a year 3 = 3-6 times a year 4 = 1-2 times a month 5 = 1-2 times a week 6 = more than 2 times a week	3.02	1.73
10. Number of books read in English before taking TOEFL 1 = none 2 = 1-5 books 3=6-10books 4 = 11-50 books 5 = more than 50 books	2.28	1.12
11. Frequency of visiting bookstores or libraries since coming to the US for pleasure reading 1 = none 2 = 1-3 times a month 3 = 1-2 times a week 4 = more than 2 times a week	1.81	.732
12. Frequency of pleasure reading in home country in native language 1 = none 2 = 1-3 times a month 3 = 1-2 times a week 4 = more than 2 times a week	2.89	.931
13. Frequency of current pleasure reading in native language 1 = none 2 =1-3 times a month 3 = 1-2 times a week 4 = more than 2 times a week	2.21	1.08

For purposes of analysis, responses to questions 7, 8, and 9 were summed to form one variable, labelled "free reading," and 12 and 13 were combined to form the variable "read in L1."

Parametic statistics were applied to the data, assuming that any violations from interval equality were not large enough to distort the results (for discussion, see Kerlinger, 1976, pp. 402-403).

TABLE 2
Intercorrelations Among Variables

Item	1	2	3	4	5	6	7	8
TOEFL								
English Study/ Home	.367							
English Study/ US	.003	-.403**						
LOR/US	.232	-.343	.496**					
TV	-.133	-.099	.019	-.104				
Free Reading	.390*	.043	.018	-.070	-.115			
Books Read	.447**	.151	.423**	.021	-.216	.714**		
Bookstores/ Libraries	.084	.007	.423**	.220	-.216	.319*	.212	
Read in L1	-.174	-.159	-.294	-.049	-.259	.048	-.125	-.056

* $p < .05$
** $p < .01$

TABLE 3
Multiple Regression Analysis: Predictors of TOEFL Scores

Predictor	b	beta	t	p
Free Reading/Books	4.996	.41	3.42	.002
English Study/Home	6.738	.48	3.73	.001
LOR/US	.937	.42	3.24	.003

r2 = .45; adjusted r2 = .41
F = 10.33, $p < .001$

All three predictors made strong showings in the multiple regression analysis.

1113

Discussion

English Study

In our analysis, years of study of English in the home country was a significant predictor. Gradman and Hanania (1991) also investigated this variable, but found no relationship to TOEFL performance. Their variable "months of intensive/special English" was significantly correlated with TOEFL scores, but correlations were modest ($r = .26$) and this variable did not survive their multivariate analysis. Previous research has shown a positive relationship between years of study and proficiency (Krashen, 1982). It seems reasonable to hypothesize, however, that study will affect performance only up to a certain level of proficiency, because of the limits of the classroom. Thus, it would not be surprising to see positive but weak and inconsistent correlations between years of study and TOEFL performance. Years of English study, however, did surprisingly well in the regression analysis, slightly outperforming the other predictors.

Length of Residence

LOR was only modestly correlated with TOEFL scores, but emerged as a substantial predictor in the multiple regression analysis. In the case of these subjects, LOR probably represented some academic experience in an English-speaking university, an important source of comprehensible input of academic language.

Free Reading in English

In Gradman and Hanania (1991), "extracurricular reading" was the strongest predictor of TOEFL scores, and free reading clearly had a very positive effect in Constantino's case history (Constantino, 1995). In this study, free reading also did well. It can be argued, however, that free reading is potentially a much stronger factor than our results indicate: Our subjects were clearly not pleasure readers, averaging fewer than 10 books read before taking the examination, and reporting reading for pleasure only three to six times per year!

We present below a separate analysis of responses to question 10 (Table 4). This analysis needs to be interpreted cautiously, because of the small number of subjects in some of the categories. Nevertheless, it suggests that free reading counts a great deal.

TABLE 4
TOEFL Scores and Number of Books Read in English

Number of Books	Mean	s.d.	Number of Subjects
None	543.10	43.03	10
1-5 Books	555.15	42.88	20
6-10 Books	583.71	63.90	7
11-50 Books	605.33	18.58	3
Over 50 Books	613.00	47.62	3

Reading in the First Language

Reading in the L1 was not a predictor of TOEFL scores. There is evidence that reading competence in the first language is a predictor of reading ability in the second language in children (Cummins, 1989). Our results may be due to the fact that our subjects, all university students, were all good readers in their primary language; there was thus insufficient variability in the sample to show a relationship between L1 and L2 competence. This explanation may also help explain Janapolous' findings of no relationship between L1 reading and L2 writing (Janapolous, 1986).

Our results have both theoretical and practical implications. They provide additional support for the Reading Hypothesis (Krashen, 1993) as well as the more general Input Hypothesis (Krashen, 1985). In addition, they confirm that preparation for the TOEFL "doesn't have to hurt" (Constantino, 1995, p. 15).

References

Constantino, R. (1995). The effect of pleasure reading: Passing the TOEFL test doesn't have to hurt. *Mosaic*, 3 (1), 15-17.

Cummins, J. (1989). Empowering minority students. Ontario, CA: California Association for Bilingual Education.

Gradman, H. and Hanania, E. (1991). Language learning background factors and ESL proficiency. *Modern Language Journal, 75*, 39-51.

Janopoulos, M. (1986). The relationship of pleasure reading and second language writing proficiency. *TESOL Quarterly, 20*, 763-768.

Kerlinger, F. (1973). *Foundations of behavioral research.* (3rd ed.). New York: Holt Rinehart.

Krashen, S. (1982). *Principals and practice in second language acquisition.* New York: Prentice Hall.

Krashen, S. (1993). *The power of reading.* Englewood, CO: Libraries Unlimited.

Authors

REBECCA CONSTANTINO Congressional Fellow, U.S. Congress, Washington, DC. Specializations: reading, English as a second language.

SY-YING LEE, National Taiwan Ocean University, 2 Pei Ning Road, Kaelung 20224, Taiwan, Republic of China. Specializations: English as a foreign language, writing, reading.

KYUNG-SOOK CHO, Department of English Education, Pusan National University of Education, 263 Koje-dong Tongnae-gu, Pusan Korea 607-071. Specializations: English as a foreign language, reading.

STEPHEN KRASHEN, Professor, School of Education, University of Southern California, Los Angles, CA 90089-0031. Specializations: language acquisition, literacy, bilingual education.

1116

REVIEWS

The Current State of Interlanguage: Studies in Honor of William Rutherford. (1995). By Lynn Eubank, Larry Selinker, and Michael Sharwood Smith, Eds. Amsterdam: John Benjamins, Pp. 293.

Reviewed by JOHN S. HEDGCOCK
Monterey Institute of International Studies

Of principal interest to researchers in second language acquisition (SLA), this volume presents a range of "state-of-the-art" essays, research summaries, and squibs that address the interlanguage (IL) construct from a predominantly formal perspective. The volume, dedicated to the career of William Rutherford, comprises chapters contributed by North American and European scholars whose contributions were selected for their connections to Rutherford's work in linguistic theory, typology, language acquisition, and/or pedagogy (e.g., Rutherford, 1975/1977, 1982, 1983, 1984a, 1984b, 1987, 1988; Rutherford & Sharwood Smith, 1987). Although several of the chapters address topics whose relevance to the study of IL (and to Rutherford's research) is questionable, this collection offers interesting and valuable contributions to empirical and theoretical debates among SLA scholars (cf. Davies, Criper, & Howatt, 1984; Selinker, 1992).

Following the volume's introduction by Lynn Eubank, Larry Selinker, and Michael Sharwood Smith, the first chapter examines the influence of primary language typological features on learner IL. In this brief essay, Peter Jordens re-examines Rutherford's (1983) developmental data on the English IL of speakers of Mandarin, a topic-prominent language. Whereas in Rutherford's analysis, two parallel developmental processes underlie the acquisition of subject-predicate strings and presentative sentences in English, Jordens maintains that the acquisition of subject prominence in L2 English is influenced directly by the prior incorporation of presentatives into the emerging IL. Rather than contradicting it, Jordens' re-interpretation complements Rutherford's initial claims by explaining why presentative sentences appear in Chinese learners' ILs before nominal and why Mandarin-speaking learners' presentative serial-verb strings never contain both a topic and a subject.

Virginia Yip and Stephen Matthews likewise address topic prominence, although from a Universal Grammar (UG) (as opposed to a typological) perspective. The authors focus specifically on pseudo-passives (e.g., "These ways can classify two types") and periphrastic topic constructions (e.g., "For Japanese speakers, they may think more positively

117

about this claim [than English speakers]") found in the L2 English compositions of Cantonese speakers. Their data on periphrastic topic constructions suggest that learners transfer L1 discourse function to L2 forms, but that underlyingly, these IL forms realize this discursive function by assigning a preposition as a case marker for the topic NP (as in "For Japanese speakers ..."). Acknowledging that Rutherford's typological interpretation of IL phenomena "help to reveal properties of interlanguage" (p. 19), Yip and Matthews maintain that IL phenomena should be characterized in terms if learners' internalized grammars, or *I-language* (Chomsky, 1986).

In contradistinction to Yip and Matthews, Susan Gass argues in her squib that learners' noticing of typological cross-linguistic differences offers a more convincing account of SLA than does the operation of abstract UG principles (cf. Gass & Ard, 1984). Based on empirical studies of the acquisition of (1) *pro*-drop, (2), verb raising, and (3) reflexive binding phenomena in a variety of target languages, Gass claims that IL grammars can gradually conform to universal clusterings of properties where learners are able to perceive "structural and/or functional relatedness" (p. 32); in cases where structural and functional relationships are not apparent on the basis of surface similarity, learning is not affected. In keeping with the spirit of much of Rutherford's work, Gass offers not only a theoretical but also a practical implication, namely, that L2 instruction will positively affect learning "in just those areas where learners are able to notice the relationships among structures" (p. 38).

The author of the volume's next entry, David Birdsong, takes readers in a decidedly practical direction by explaining how the game "Master Mind" can assist students of SLA to understand abstract, theory-dependent constructs such as learnability, pre-emption, and domain specificity. Although such a contribution in a volume on IL may seem to be somewhat of a thematic oddity, Birdsong's chapter connects very neatly to Rutherford's contributions to the training of future SLA researchers (cf. Rutherford, 1996). Birdsong presents a "tutorial" exemplifying the ways in which this strategic game engages its players in narrowing broad hypotheses, broadening overly narrow hypotheses, and rejecting inaccurate hypotheses.

Returning to the theoretical domain, Ellen Bialystok highlights the dangers of overgeneralization in SLA by pointing out the errors caused by theory-building that overlooks the "accumulated wisdom of the specialists" (p. 56). Specifically, Bialystok's essay makes a strong case for the centrality of linguistics in the search for a comprehensive IL theory—an interesting twist for a well-known cognitivist. Bialystok's view that

1118

linguistic structure constitutes a powerful determinant of cognitive processing supports Rutherford's recent arguments that any and all theorizing about language learning must be solidly grounded in the structural properties of language and learners' knowledge of it (cf. Rutherford, 1994, 1996).

Lydia White's essay complements Bialystok's by amplifying a persistent appeal made by UG proponents, namely, that a theory of language acquisition *must* be informed by (and derived directly from) a theory of language (White, 1989, 1994). Focusing specifically on Chomsky's "Minimalist" program (Chomsky, 1991, 1993, 1994), White holds that productive empirical research in SLA is not only stimulated by, but dependent on, advances in linguistic theory. Although caution is necessary when applying new theories to empirical findings, White makes the controversial assertion that "adjustments have to be made in the area of L2 acquisition theories...when a linguistic theory is discarded or outgrown" (p. 68).

Turning to a more specific empirical problem in IL research, Susanne Carroll examines aspects of verbal linguistic feedback that make interactional L2 data usable for learners. Drawing on Sperber and Wilson's (1986) theory of relevance, Carroll maintains that, because negative input (i.e., information concerning allowable strings, forms, patterns, and constraints in the L2) requires metalinguistic mastery on the part of learners, it is "comparatively useless" (p. 85) in the L2 development of most learners. Like Gass, Carroll thus holds that learners must perceive formal aspects of the L2 grammar before they can successfully incorporate this information into their ILs.

One of the few wholly data-based chapters in this volume is that of India Plough, who tests the prediction that inductive inferencing plays a positive role in adult L2 (specifically, foreign language) learning. Plough adopts a view in which inductive inferencing results in "probably true" hypotheses about L2 patterns derived from a multi-stage comprehension process. Her study of university-level anglophone learners of French is one of the first serious attempts to define indirect negative data and to test for its usefulness in L2 learning. Although her investigation does not offer longitudinal findings, Plough's experimental results show that inferencing input can be a predictor of L2 learning outcomes and possibly a correlate of language aptitude.

Attacking the overlooked variable of L2 phonological attainment, Martha Young-Scholten offers a critical appraisal of research suggesting that incomplete mastery of L2 phonological patterns supports the view that UG is not fully accessible to adult L2 learners. Young-Scholten proposes an alternative, transfer-based interpretation in which she makes

the unique and initially counter-intuitive claim that orthographic input may, in fact, be detrimental to the development of target-like L2 phonology. This can occur, she argues, when learners are not yet psycholinguistically ready to incorporate orthographic knowledge into their IL systems (cf. Corder, 1981).

Harald Clahsen's chapter, which summarizes select findings from the ZISA (*Zweitsprachenerwerb Italienischer und Spanischer Arbeiter*, "Acquisition of Second Language by Italian and Spanish Workers") project (Clahsen, 1984), is situated within the ongoing controversy surrounding the psycholinguistic representation of regular and irregular morphology among adult language learners whose ILs appear to have fossilized. Focusing specifically on the system of noun plurals in L2 German, Clahsen's results indicate that learners who "overregularize" inflectional forms have two qualitatively distinct types of plural forms in their IL grammars: irregular forms that only occur in compounds and overregularized default plurals that do not occur in plurals. These outcomes support a dual-mechanism model of psycholinguistic representation in which rule-based patterns coexist with rote-based knowledge systems (Pinker & Prince, 1988).

Sascha Felix returns to the UG domain by addressing one of the central problems of the principles and parameters paradigm: Adult learners' access to UG principles and the likelihood that parameters can be re-set in L2 acquisition. Felix presents a meticulous analysis of Schachter's (1996) strong claim that UG is in no way accessible to adult L2 learners. Indeed, Felix demonstrates that Schachter's recent investigation of the Incompleteness Hypothesis in fact provides evidence that L2 learners have UG information available to them thanks to a type of knowledge that instantiates parameterized principles. Felix argues forcefully that the Incompleteness Hypothesis is untenable "because it crucially relies on the assumption that natural languages may instantiate only a subset of UG principles, an assumption which does not seem to be compatible with what we currently know about [UG]" (p. 149).

Based on a series of empirical studies, Antonella Sorace's contribution examines in some depth the acquisition of the unaccusative-unergative distinction in a set of intransitive verbs in L2 Italian. Her analysis of the results of acceptability judgment tests is unique in that she explains learner IL development in terms of linking rules that bind lexical-semantic representations to argument structure (cf. Baker, 1983). Specifically, her findings suggest that unaccusative-unergative distinctions develop gradually, though often asymmetrically, along developmental paths governed by lexical-semantic hierarchies that subdivide intransitive verbs. This work expands on existing findings that reveal

1120

the close interrelationships between lexical and syntactic knowledge in adult L2 learning.

Maria Beck, Bonnie Schwartz, and Lynn Eubank contribute to another ongoing debate among UG researchers in their three-part paper titled "Data, evidence, and rules." Following a survey of input data types that are eligible for integration into IL grammar, they assert that not all available L2 data are actually used by learners. They refer to that subset of linguistic data that are actually perceived and used by learners as "evidence." Relying largely on exemplification, the authors strongly hint that explicit evidence and "instructional therapy" may actually play a minor role in adult L2 acquisition because it "cannot engage UG" (p. 191). The outcome of instruction, they claim, amounts to "Learned Linguistic Knowledge" (p. 192)—an interpretation that competes with the view that UG parameters can be activated by awareness-raising and input enhancement procedures (e.g., Rutherford, 1987; Sharwood Smith, 1991).

Helmut Zobl's squib again directs the reader's attention to formal L2 properties, in this case, lexical vs. pronominal subjects in the English IL of adult francophone learners and the German IL of an adult Turkish-speaking learner. Zobl asks why, despite abundant grammatical L2 data, structural government for nominative case-marking in the respective target languages does not emerge in learner IL. Instead, IL evinces a pattern of incorporation (Baker, 1988) in which pronominals but not lexical NPs appear in target-like L2 structures. Zobl holds that structural government for nominative-case-marking is marked for L2 learners, thereby requiring checking across finite IP boundaries. Zobl thus concludes that incorporation serves as a default procedure for L2 learners, such that "recourse to default options for dealing with computational complexity may be an acquisitional universal" (p. 202).

Returning more directly to the construct of interlingual transfer, Gita Martohardjono and Suzanne Flynn summarize data showing that Japanese-, Chinese-, and Spanish-speaking L2 learners of English may completely ignore available L1-based data as they acquire control structures, in terms of both the lexicon and surface-level grammar. Their empirical results lead them to conclude that these learners "resort to a principle of UG, minimality." Based on this bold conclusion, the authors further contend that "L2 learners do not assume that the L2 is like the L1" (p. 216) and that the Transfer Hypothesis does not hold. Unfortunately, the authors fail to qualify these controversial and disputable interpretations.

Eric Kellerman revisits still-unresolved questions related to fossilization in SLA. Invoking such variables as age, cross-linguistic

1121

transfer, typology, exposure, metalinguistic knowledge, and comparative linguistic difficulty, the author re-examines the work of Johnson and Newport (1989), which he claims is highly problematic for methodological and theoretical reasons. Crucially, this essay details the reasons for which responses to grammaticality judgment tasks fail to characterize IL knowledge at a given point in time, let alone longitudinally. Kellerman concludes with a compelling appeal to undertake extended and careful empirical research on the variable difficulty of particular linguistic patterns and to approach the Critical Period Hypothesis with a healthy skepticism.

Jean-Marc Dewaele's squib, which proposes explicit procedures for distinguishing between different styles in IL speech, extends the domain of IL somewhat by making a case for the role of context in SLA research. The author suggests that attention to speech dimensions (*viz.*, formal/informal, explicit/implicit) is important in IL studies not only because of the insights into formal variation provided by these constructs, but also because of the differential cognitive demands associated with these multiple dimensions.

Of questionable relevance to the major themes of this volume is the chapter by Susana Blanco-Iglesias, Joaquina Broner, and Elaine Tarone. In this piece, the authors seem to address purposes whose connection to IL in general and to Rutherford's work in particular is difficult to discern. For example, considerable space is devoted to an evaluative description of note-taking as a data collection tool in classroom research. This methodological discussion is somewhat of a red herring, given the chapter's other major focus, changes in language choice and vernacular language use in kindergarten through grade five Spanish immersion classrooms. Readers might be disappointed to read the authors' rather obvious conclusion that older anglophone children revert to using English because "they need to use a vernacular 'kid-speak' with each other [*sic*], and they do not have such a vernacular in the L2Ñonly an L1 vernacular" (p. 252).

Marjorie Perlman Lorch also addresses the topic of variation, although with a particular emphasis on bilingual aphasia and neurolinguistic research. In her predominantly context-dependent view, Lorch highlights the importance of environmental cues in learners' establishment of context-specific fossilized IL forms. Based on an overview of experimental research on aphasia and language switching (e.g., Paradis, 1989), Lorch challenges traditional modular views of language, arguing that IL talk that is contextually sensitive must depend not only on evolving IL phonology, lexicon, morphology, and syntax, but also on

1122

knowledge and cognitive skills that lie outside the so-called language modules.

Like the Blanco-Iglesias et al. chapter described above and the paper by Ross that follows it, Batia Laufer's paper on L2 lexical acquisition bears only a sketchy connection to Rutherford's research and to the major issues addressed elsewhere in this book. In addition to summarizing several investigations of methods aimed at characterizing the size of L2 learners' lexicons, Laufer describes "Beyond 2000," a computerized metric whose accuracy outstrips that of its predecessors. Her central theses are that (1) growth in vocabulary size is the most appropriate indicator of lexical acquisition, and (2) the "Beyond 2000" measure is a valuable and reliable means of undertaking future research on L2 vocabulary development.

The final paper in the volume, authored by Háj Ross, promises a "first look" at *paths*, defined as "those macro-constituents which specify the route through spaces of various sorts which is traversed by the Theme of a sentence" (p. 272; cf. Gruber, 1976). The author does, indeed, provide a first look at paths (later described as "constituents") and their parts (called "legs"). Ross's cross-linguistic analysis compares end and medial "legs" in English to their structural equivalents in both Brazilian Portuguese and German; he concludes that, in all three languages, end-prepositions have more degrees of freedom than do medial prepositions. Invoking markedness theory, Ross proposes that if the expression of certain legs is less marked than adpositional expressions, then a learner whose L1 exhibits marked patterns and who learns an L2 with unmarked patterns should find L2 acquisition easier than the learner for whom the situation is reversed. Whereas the author proposes an intriguing approach to syntactic and thematic relations, the relevance of this work to SLA remains speculative and unclear.

As one might expect in a volume of this type, the papers in *The Current State of Interlanguage: Studies in Honor of William Rutherford* vary in scope, depth of analysis, and thematic coherence. Notwithstanding its flaws (i.e., a discernible principles and parameters bias, "empirical" papers that contain very little data, chapters whose contents go somewhat far afield of the book's topical focus, incomplete and/or erroneous bibliographical entries, and a noticeable number of typos), the majority of its chapters offer provocative treatments of theoretical, empirical, and methodological issues of ongoing concern to SLA researchers of various orientations. Taken together, the papers constitute a worthy tribute to a researcher whose influence as a teacher, scholar, and advocate of the field deserves recognition and praise.

1123

References

Baker, M. (1983). Objects, themes, and lexical rules in Italian. In L. Levin, M. Rappaport, & A. Zaenen (Eds.). *Papers in lexical-functional grammar.*

Baker, M. (1988). *Incorporation: A theory of grammatical function changing.* Chicago: Chicago University Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* New York: Præger.

Chomsky, N. (1991). Some notes on economy of derivation and representation. In N. Hornstein & D. Lightfoot (Eds.), *Principles and parameters in comparative grammar* (pp. 417-454).

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 1-52). Cambridge, MA: MIT Press.

Chomsky, N. (1994). Bare phrase structure. *MIT Occasional Papers in Linguistics, 5.*

Clahsen, H. (1984). The acquisition of German word order: A test case for cognitive approaches to L2 development. In R. Andersen (Ed.), *Second languages: A crosslinguistic perspective* (pp. 219-242).

Corder, S. P. (1981). *Error analysis and interlanguage.* Oxford: Oxford University Press.

Davies, A., Criper, C., & Howatt, T. (Eds.). (1984). *Interlanguage.* Edinburgh: Edinburgh University Press.

Gass, S., & Ard, J. (1984). Second language acquisition and the ontology of language universals. In W. R. Rutherford (Ed.), *Typological universals and second language acquisition.* Amsterdam: John Benjamins.

Gruber, J. (1976). *Lexical structures in syntax and semantics.* Amsterdam: North-Holland Publishing Company.

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of ESL. *Cognitive Psychology, 21,* 60-99.

Paradis, M. (1989). Bilingual and polyglot aphasia. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 117-140). Amsterdam: Elsevier.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 73-193.

1124

Rutherford, W. R. (1975/1977). *Modern English* (Revised and expanded edition). New York: Harcourt, Brace, Jovanovich.

Rutherford, W. R. (1982). Markedness in second language acquisition. *Language Learning, 32,* 85-108.

Rutherford, W. R. (1983). Language typology and language transfer. In S. Gass & L. Selinker, (Eds.), *Language transfer in language learning* (pp. 358-370). Ann Arbor, MI: University of Michigan Press.

Rutherford, W. R. (Ed.). (1984a). *Typological universals and second language acquisition.* Amsterdam: John Benjamins.

Rutherford, W. R. (1984b). Description and explanation in interlanguage syntax: State of the art. *Language Learning, 34,* 127-155.

Rutherford, W. R. (1987). *Second-language grammar: Learning and teaching.* London: Longman.

Rutherford, W. R. (1988). Grammatical theory and L2 acquisition: A brief overview. In S. Flynn & W. O'Neil (Eds.), *Linguistic theory in second language acquisition* (pp. 1-25). Dordrecht, The Netherlands: Reidel.

Rutherford, W. R. (1994, March). *Teaching linguistics and the acquisition research literature.* Paper presented at the Annual Meeting of the American Association for Applied Linguistics, Baltimore, MD.

Rutherford, W. R. (1996a, October). *Crossing boundaries and maintaining perspective.* Plenary address delivered at the Second Language Research Forum, Tucson, AZ.

Rutherford, W. R. (1996b). *Workbook in the structure of English: Linguistic principles and language acquisition.*

Rutherford, W. R., & Sharwood Smith, M. (Eds.). (1987). *Grammar and second language teaching: A book of readings.* New York: Newbury House.

Selinker, L. (1992). *Rediscovering interlanguage.* London: Longman.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7,* 118-132.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.

White, L. (1989). *Universal grammar and second language acquisition.* Amsterdam: John Benjamins.

White, L. (1994, October). *Can UG be manipulated in second language acquisition?* Plenary address delivered at the Second Language Research Forum, Montréal.

1125

Writing and Literacy in Chinese, Korean, and Japanese. (1996). By Insup Taylor with M. Martin Taylor. Philadelphia, PA: John Benjamins.

Reviewed by JANE KUO and SUGURU AKUTSU
The American Graduate School of International Management

Writing and Literacy in Chinese, Korean, and Japanese by Insup and M. Martin Taylor comes at a time when the world is hungering to learn more about the people and the cultures that have created the "Asian economic miracle." Since people and their culture can be understood through the study of their language, this book provides an alternative means of gaining insight into three Asian cultures: Chinese, Korean, and Japanese. The authors' knowledge of these languages is profound, making possible a thorough, interesting, and enlightening examination of the languages of Asia. This investigation then leads to posing and answering broad questions about the advantages and disadvantages of each of the languages with regard to the attainment of literacy and technical thinking in Asia.

The book is organized by the chronological development of the three languages: Chinese, Korean, and Japanese. The authors begin with an introductory chapter discussing the basic elements that make these three languages distinct from others. In this chapter, readers are also informed about the organization of the book and the inclusion of particular issues with regard to writing and literacy in each part. Each section begins with a general history of the three Asian languages, and then moves on to discuss and draw conclusions on each language separately.

The sections on linguistic units and terms such as morphemes and phonemes are especially interesting. They deepen the reader's comprehension of some of the challenges that face both students and teachers of each language. Scholars, educators, and students of East Asian Studies will find useful information about the influence of written language on the cultural development. This is particularly true of the section titled "History of Education and Literacy," which discusses the influence of Confucianism and the ancient civil service examination system on contemporary Asian cultures. Educators working to bridge linguistic gaps between East and West will find that this book allows students to broaden their learning experiences. The bibliography provided at the end of each section is most useful to the reader. It not only lists references, but also gives a brief description of the content of each source, allowing the reader to find further information on a topic of interest.

In each part (Chinese, Korean, and Japanese), the authors set out to address six fundamental questions:

> 1. How do the Chinese, Korean, and Japanese writing systems differ?
> 2. How are these differences reflected in each country's literacy rate?
> 3. Are Chinese characters a hindrance to high levels of literacy?
> 4. Do the writing systems in these nations provide a suitable foundation for the pursuit of innovative and technical thinking?
> 5. Are these writing systems appropriate in the information age?
> 6. What alterations, if any, should or could be made?

Indeed, the authors specifically cite some of these questions in their preface and throughout the text. However, in the main body of this text the authors devote a great deal of attention to other issues, leaving these central questions to the summaries and general conclusions. For example, the text goes into great detail regarding the phonetic pronunciation of each language. Without previous formal training, though, the reader has no way to comprehend these complex phonetic systems. Therefore, it is difficult to agree with the authors' contention that this book is written for the general public.

Part one includes ten chapters on the written Chinese that clearly demonstrate the authors' command of the language. The structure of the language is described in detail. In fact, the authors go too deeply into these details. Although the structure of the language, the origin and classification of characters, and phonetic systems are all important elements of the text, they should not be focused on at the expense of broader issues. The authors discuss such topics as the creation of characters, the evolution of script styles, and the six principles of categorizing Chinese characters.

Consequently, readers gain insight about how Chinese served as the forerunner of the written forms of Korean and Japanese as well as about characters that tell stories, and so forth. For those interested in the "vertical" (historical development within a language) and "horizontal" (mutual influence between languages) aspects of these three languages, and the impact written characters have had on the education systems and literacy in these countries, this book is invaluable. In addition, a substantial chapter on the reform of spoken and written Chinese in the

1127

People's Republic of China explains how the characters used in mainland China differ from those used in Taiwan. The general public will be fascinated by Mao Zedong's radical views on the reform of written Chinese.

The authors discuss the peculiarities of the Korean language in part two of the study. The bulk of this section effectively deals with the central questions raised by the authors, particularly in chapters 13 through 16. The reader learns about the unique characteristics of the Korean writing system. The ability of this writing system to seemingly combine a phoneme system with a morpheme system in what are called "syllable blocks" is fascinating. One interesting tool that helps the reader answer the central questions suggested by the title is the graph of the declining use of *Hancha* over the years. In the section "Rational *Hancha*.Use in the Present," the reader learns of the decline of the use of *Hancha*. In the following section, "Misguided Attempts to Abolish Hancha," the reader learns why the trend has leveled out—that *Hancha* is still a valuable communication tool even if mixed in with the *Han' gul* writing system. Points such as these are valuable to the reader in trying to understand such complex writing systems and the challenges inherent in attempts to achieve high literacy rates in Asian nations. In addition, the authors' comparisons of the strengths and weaknesses of the *Hancha* and *Han' gul* systems are thoroughly presented and very informative. Yet, even in these sections, the authors' presentation demonstrates their failure to focus on a particular audience. Even readers with a knowledge of Asian linguistics may be unable to develop a clear understanding of the authors' descriptions. We believe that in certain circumstances the authors assume the reader possesses technical linguistic knowledge that the general public or an average student simply does not have. For instance, in the section headed "Instruction in *Han' gul* Spelling" the authors explain two rules for spelling. After reading this section, we were disappointed that we were still quite unable to understand the *Han' gul* examples written throughout the remainder of the text. If the authors had included some simple examples of building a syllable block we would grasp additional new points as the authors bring them forth. Thus we might have been able to decipher the examples in succeeding chapters and sections.

In the third and final section of the book the authors attempt to answer the central questions in terms of the Japanese language. As in the previous sections, the authors' discussion progresses from a basic linguistic background toward a detailed examination of the writing and literacy of the Japanese. The introduction of the speech element called *mora*, explaining the rhythmic sound of Japanese speech, should be very

1128

interesting to some readers. Additionally, the demonstration of *moras* as the unit of measure in *Haiku* poetry will be enlightening to *Haiku* lovers. Chapters 18 through 22 involve the reader in an analysis of the writing systems of the Japanese language. These chapters are useful in preparation for the more direct discussion of the central questions raised by the text. However, at this point in the book, the reader could easily have grown weary of the discussion of *Kanji* (Chinese characters). In previous sections of the book similar points were made about *Hanzi* and *Hancha*, and few, if any, new ideas are brought forth in the analysis of *Kanji* in Chapter 18. Writing and literacy become the central topics in chapters 22 through 24, in which the authors examine learning processes, educational systems, and educational philosophy. In these sections the authors also discuss the success of Japanese primary and secondary education and identify problems of universities and higher-education.

In both the second and third parts of the text, the examination of the Korean and Japanese writing systems clearly demonstrates the conflict between phoneme-based and morpheme-based systems. The authors properly address whether or not the *Hancha* or *Kanji* should be kept or abolished. The challenges presented by the information age are more directly addressed in this section as well. From these sections the reader gains a more thorough understanding of the Korean and Japanese written languages, but gains few ideas addressing the central issue of promoting literacy in Asia.

In summary, we believe that the authors went too far in trying to teach the languages to the readers. Before introducing the writing systems, the authors go into too much depth by introducing the speech sounds, words, and sentences of the languages. Without real instruction, it is difficult to grasp the speech sounds from mere description. Thus as a result, the description of the speech sounds, particularly in chapter 13 of part two is unnecessary. Therefore, we do believe that due to the intense focus on East Asian linguistics, a reader who does not know of at least one of the languages might have difficulty understanding some of the more complex nuances of these languages.

In addition, the bulk of this text describes in great detail the complex writing system of these languages. Although such background is essential for any worthwhile discussion of writing and literacy, we found some information fragmented, difficult to understand, and excessive. The authors' organizational approach was to divide the text using national-linguistic lines. This is a natural approach and it works well in many circumstances, but we found ourselves frequently turning back to previous sections to compare and contrast the different languages.

1129

Our approach would divide the book into sections dedicated to elements of the written languages. The first section would discuss the Chinese writing system in which the history, development, advantages, disadvantages, complexity, adoption and rejection of characters would be examined in terms of all three countries. The following sections would compare and contrast the phonetic sytems used, educational systems employed and literacy rates achieved by each nation. A final section would be devoted to the authors' conclusions on advantageous elements of each written language. Organized in this way, the text could achieve a more direct comparison of the three languages and keep the reader engaged in the examination.

Literate Apprenticeships: The Emergence of Language and Literacy in the Preschool Years. (1996). By Kenneth Reeder, John Shapiro, Rita Watson, and Hillel Goelman. Norwood, NJ: Ablex.

Reviewed by LIA PLAKANS
Iowa State University

This volume, in the series *Advances in Discourse Processes*, deals with the emergence of literacy in preschool-aged children. Four researchers—Kenneth Reeder, Jon Shapiro, Rita Watson, and Hillel Goelman—collected data for three years at the Child Study Centre at the University of British Columbia. The results of their efforts answer the general question, "What are the relationships amongst children's earliest literate environments, their emerging knowledge about literacy, and some specific aspects of their growing communicative competence" (p.14). In the introduction Anthony D. Pellegrini places the study within the current research on emergent literacy and gives the reader insight into the theoretical orientation of the researchers (developmental psycholinguistics). Then, the researchers themselves explain their goals and further ground their study in the current literature. One important goal is to provide information about emergent literacy to preschool teachers and parents. The subsequent four chapters describe the four studies conducted by the researcher, and a final chapter summarizes the findings and suggests what can be learned from them.

In the first study, Jon Shapiro measured the acquisition of literacy of a group of children and investigated the relation between this development and the children's home literacy environment. By administering a survey called the *Home Literacy Environment Index (HLEI)* to the parents of the children, Shapiro established which children were from

1130

environments with a higher or lower literate biases. Repeated testing of the childrens' knowledge and metaliterate awareness was done through a battery of tests given annually. From these data, Shapiro drew several conclusions: (1) metaliterate awareness is developmental, (2) metaliterate awareness has many layers and is not a single unit, and (3) there is a relationship between home literate environments and the emergence of metaliterate awareness.

The second study, by Kenneth Reeder, investigated whether children from higher literacy environments connect pragmatic understanding to an utterance based on its linguistics cues, while those from a lower literacy environment tend to use context as a clue to meaning. The results of this study showed a possible developmental trend in speech act comprehension. Reeder found that at the younger age, high literacy children focused more on linguistic cues, but, by the third year of the study, the same group had switched to using both context and linguistic cues. On the other hand, the low literacy children began focusing on context cues, and in the third year had shifted to more linguistic cues. According to Reeder this reveals a development process in which the high literate group may have focused on context cues earlier than the beginning of the study, while the low literacy group would proceed to the level of using both context and linguistic cues to determine pragmatics after the study .

The third study, by Rita Watson, explores the literate discourse of preschool-age children to discover whether certain forms of discourse which are experienced by children in early literacy activity can predict future metaliterate skills. After collecting data from parent-child interactions during book-reading sessions, Watson coded their discourse for certain features: questions and responses, labeling, propositional attitudes, and the textuality (discourse specifically related to the text). Through analysis of these categories and the results of the assessment of the children's metaliteracy and literacy related skills (e.g., word and label identification), Watson found that a correlation exists between early literate discourse and metaliterate skills of the children. The correlation, however, fluctuated over the three years. The results also suggested that negative parental responses and direct teaching were less effective in increasing metaliterate skills than parental interaction, which is less controlled and contains more open responses to children's utterances.

In the fourth study, Hillel Goelman studied the oral discourse of parent-children interactions in terms of two issues: (1) whether certain features of oral discourse already established by other researchers to have a literate bias (psychological verbs, cohesion, cognitive demands) can be found in his sample, and (2) what the relationship is between

1131

such discourse features and other kinds of literacy. The assumption behind this study is that literate does not exist only in written texts but also in oral discourse through its organizational pattern and word choice. By collecting data in parent-child play sessions and coding the resulting discourse, the author determined that the features were indeed evident in discourse during play sessions and that there does seem to be a correlation between such features and three aspects of early literacy: understanding of symbolic concepts, print concepts, and story concepts.

This collaborative project adds a great deal to our understanding of emergent literacy. The longitudinally collected data results in very rich and deep source for information. A few limitations in the study are acknowledged by the authors, however. One of these concerns is the population studied. Because the children were all in the same university childcare center, the socio-economic and educational level of their families did not vary greatly. Therefore, readers should be cautious in generalizing the results. Also, the measure *Home Literacy Environment Index*, used in the study to determine the level of literacy in the children's home environment, might be limited in application. As the authors mention, some information sought by this survey could be missing or misinterpreted. Interviews and home-observations would possibly have given a more accurate picture of the home environment. Nonetheless, this project generated much useful information about children's early literacy environments, their emerging knowledge about literacy, and their emergent communicative competence. There is much in these studies that could be useful to preschool teachers and parents, as well as to others interested in the development of children's literacy .

1132

General Information

Calendar of Events*

1997

20 July-1 August, *Taos Institute for Language Teachers (French, German, Russian, Spanish),* East Lansing. Information George F. Peters, CLEAR, A-126 Wells Hall, Michigan State University, East Lansing, MI 48824-1027; (517) 432-2286, Fax (517) 432-0473, Email [clear@pilot.msu.edu].

5-7 August, *Problems of Teaching Modern Hebrew*, Jerusalem. Information Ben-Zion Fischler, Council on the Teaching of Hebrew, P.O.B. 7413, Jerusalem 91073, Israel.

12-16 August, *IALL/Language Laboratory Association of Japan*, Victoria, BC, Canada. Information (604) 721-8294, Fax (604) 721-8778, Email [fleatIII@call.uvic.ca], URL [http://web.uvic.ca/hrd/fleat3].

20-22 August, *OPI Training Workshop (French, German, Spanish)*, Lisle, IL. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

4 September, *Association of Literary Semantics*, Freiburg. Information Monika Fludernik, English Department, University of Freiburg, D-79085, Germany.

24-27 September, *OPI Training Workshop (English-LCT, French, German, Japanese, Russian, Spanish)*, Madison, WI. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

25-28 September, *German Studies Association,* Washington, D.C. Information German Studies Association, Arizona State University, Tempe, AZ 85287-3204.

4-5 October, *Carolinas Symposium on British Studies,* Augusta. Information William S. Brockington, Department of History, University of South Carolina at Aiken, Aiken, SC 29801.

7-11 October, *Foreign Language Association of North Carolina,* location to be announced. Information Mary Lynn Redmond, 6 Sun Oak Court, Greensboro, NC 27410; Fax (910) 759-4591, Email [redmond@wfu.edu].

17-20 November, *OPI Training Workshop (Chinese, French, German, Italian, Japanese, Russian, Spanish),* Nashville. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

18-20 November, *Educational Exchange,* Barcelona. Information CIEE, 205 East 42nd St., New York, NY 10017-5706; (212) 822-2699, Email [conference@ciee.org].

*Courtesy of *The Modern Language Journal* (University of Wisconsin)

19-20 November, *National Association of District Supervisors of Foreign Languages*, Nashville. Information Sharon M. Watts, 607 S. 123 St., Omaha, NE 68154; (402) 557-2440, Email [swatts@ops.esu19.k12.ne.us]

20-22 November, *European Association for International Education,* Barcelona. Information Conference Director, EAIE, Van Diemenstraat 344, 1013 CR Amsterdam, The Netherlands; Fax (31) (20) 620-9406; Email [eaie@eaie.nl].

21-23 November, *American Council on the Teaching of Foreign Lan guages,* Nashville. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

21-23 November, *American Association of Teachers of French,* Nashville. Information AATF,57 E.Armory Dr., Champaign, IL 61820; (217) 333-2842, Fax (217) 333-5850, Email [fmajatf@vmd.cso.uiuc.edu].

21-23 November, *American Association of Teachers of Spanish & Portuguese,* Nashville. Information AATSP, Frasier Hall #8, University of Northern Colorado, Greeley, CO 80639; (970) 351-1090, Fax (970) 351-1095, Email [seelsandste@bentley.univnorthco.edu].

21-23 November, *American Association of Teachers of German,* Nashville. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America,* Toronto. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, Toronto. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520) 885-2663, Email [76703.2063@compuserve.com].

1998

26-28 February, *Southern Conference on Language Teaching with Foreign Language Association of Georgia,* Savannah. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912) 333-7358, Fax (912) 333-7389, Email [lbradley@grits.valdosta.peachnet.edu].

14-17 March, *American Association of Applied Linguistics,* Seattle. Information AAAL, 7630 West 145th Street, Suite 202, Apple Valley, MN 55124; (612) 953-0805, Fax (612) 891-1800, Email [howe@mr.net].

17-21 March, *Teachers of English to Speakers of Other Languages,* Seattle. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703) 836-0774, Fax (703) 836-

7864, Email [conv@tesol.edu].

26-29 March, *Central States Conference on the Teaching of Foreign Languages,* Milwaukee. Information CSCTFL, Rosalie Cheatham, University of Arkansas-Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; (501) 569-8159, Fax (501) 569-3220, Email [rmcheatham@ualr.edu].

16-19 April, *Northeast Conference on the Teaching of Foreign Languages,* New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu].

24-26 April, *Pacific Northwest Council for Languages,* Boise. Information PNCFL, Foreign Languages and Literatures, Oregon State University, 210 Kidder Hall, Corvallis, OR 97331-4603; Fax (541) 737-3563, Email [verzascr@cla.orst.edu].

23-26 July, *American Association of Teachers of French*, Montreal. Information Jayne Abrate, Executive Director, AATF, Department of Foreign Languages and Literatures, Southern Illinois University, Carbondale, IL 62901.

September dates to be announced, *International Congress of Slavists*, Cracow. Information Robert Rothstein, Department of Slavic Languages, Herter Hall, University of Massachusetts, Amherst, MA 01003-3940.

27-31 October, *Foreign Language Association of North Carolina,* location to be announced. Information Mary Lynn Redmond, 6 Sun Oak Court, Greensboro, NC 27410; Fax (910) 759-4591, Email [redmond@wfu.edu].

20-22 November, *American Council on the Teaching of Foreign Languages,* Chicago. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

20-22 November, *American Association of Teachers of German,* Chicago. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America,* location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com].

1135

1999

8-14 March, *Teachers of English to Speakers of Other Languages*, New York. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703) 836-0774, Fax (703) 836-7864, Email [conv@tesol.edu].

11-13 March, *Southern Conference on Language Teaching*, Virginia Beach. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912)333-7358, Fax (912)333-7389, Email [lbradley@grits.valdosta.peachnet.edu].

7-10 April, *Pacific Northwest Council for Languages*, Tacoma. Information PNCFL, Foreign Languages and Literatures, Oregon State University, 210 Kidder Hall, Corvallis, OR 97331-4603; Fax (541) 737-3563, Email [verzascr@cla.orst.edu].

8-11 April, Northeast Conference on the Teaching of Foreign Languages, New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu].

19-21 November, *American Council on the Teaching of Foreign Languages*, Dallas. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

19-21 November, *American Association of Teachers of German*, Dallas. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609)795-9398, Email [73740.3231@compuserve.com].

7-30 December, *Modern Language Association of America*, 2 location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com].

2000

21-25 March, *Teachers of English to Speakers of Other Languages*, Vancouver. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751;(703) 836-0774, Fax (703) 836-7864, Email [conv@tesol.edu].

17-19 November, *American Council on the Teaching of Foreign Languages*, Boston. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

1136

17-19 November, *American Association of Teachers of German,* Boston. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America,* location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *American Association of Teachers of Slavic & E. Euro pean Languages,* location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com].

2001

16-18 November, *American Council on the Teaching of Foreign Languages,* Washington, D.C. Information ACTFL, 6 Executive Plaza, Yon kers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

16-18 November, *American Association of Teachers of German,* Washington, D.C. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America,* location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *American Association of Teachers of Slavic & E. Euro pean Languages,* location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com].

1137

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (*ALL*) is to increase and promote professional communication within the Defense Foreign Language Program and academic communities. *ALL* publishes research and review articles, research reports as well as reviews on adult language learning for functional purposes. The editor encourages the submission of research and review manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Research Article

Divide your manuscript into the following sections:

- Abstract
  - Introduction
    - Method
      - Results
        - Discussion
          - Conclusion
            - Appendices
              - References
                - Author

Introduction

In a few paragraphs state the purpose of the study and relate it to the hypothesis and the experimental design. Point out the theoretical implications of the study and relate them to previous work in the area.

Next, under the subsection *Literature Review*, discuss work that had a direct impact on your study. Cite only research pertinent to a specific issue and avoid references with only tangential or general significance. Emphasize pertinent findings and relevant methodological issues. Provide the logical continuity between previous and present work. Whenever appropriate, treat controversial issues fairly. You may state that certain studies support one conclusion and others challenge or contradict it.

Method

Describe how you conducted the study. Give a brief synopsis of the method. Next develop the subsections pertaining to the *participants*, the *materials*, and the *procedure*.

Participants. Identify the number and type of participants. Specify how they were selected and how many participated in each experiment.

Identify the number of experiment dropouts and the reasons they did not continue.

Provide major demographic characteristic such as age, sex, geographic location, and institutional affiliation. Indicate that the treatment of participants was in accordance with the ethical standard of the APA (Principle 9).

Apparatus. Describe briefly the materials used and their function in the experiment.

Procedure. Describe each step in the conduct of the research. Include the instructions to the participants, the formation of the groups, and the specific experimental manipulations.

Results

First state the results. Next describe them in sufficient detail to justify the findings. Mention all relevant results, including those that run counter to the hypothesis.

Tables and figures. Prepare tables to present exact values. Use tables sparingly. Sometimes you can present data more efficiently in a few sentences than in a table. Avoid developing tables for information already presented in other places. Prepare figures to illustrate key interactions, major interdependencies, and general comparisons. Indicate to the reader what to look for in tables and figures.

Discussion

Express your support or nonsupport for the original hypothesis. Next examine, interpret, and qualify the results and draw inferences from them. Do not repeat old statements. Create new statements that further contribute to your position and to readers understanding of it.

1139

Conclusion

Succinctly describe the contribution of the study to the field. State how it has helped to resolve the original problem. Identify conclusions and theoretical implications that can be drawn from your study.

Review Article

It should describe, discuss, and evaluate several publications that fall into a topical category in foreign language education. The relative significance of the publications in the context of teaching realms should be pointed out. A review article should be 15 to 20 double-spaced pages.

Research Report

It should present and discuss data obtained in a research project in the area of foreign language education. A research report should be 5 to 10 double-spaced pages.

Review

Reviews of textbooks, scholarly works on foreign language education, dictionaries, tests, computer software, video tapes, and other non-print materials will be considered for publication. Both positive and negative aspects of the work(s) being considered should be pointed out. The review should give a clear but brief statement of the work's content and a critical assessment of its contribution to the profession. Quotations should be kept short. Reviews that are merely descriptive will not be accepted for publication. The length of the manuscript should be three to five double-spaced pages.

Submission of Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-*ALL*
ATTN: Editor (Dr. L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts are accepted for consideration with the understanding that they are original material and are not being considered for publication elsewhere.

1140

Specifications for Manuscripts

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784. Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run from 30 to 50 pages.

All manuscripts should be sent with an author identification, an abstract, a list of references, and, if necessary, notes.

Author Identification. The title of the article and the author's name should be typed on a separate page to ensure anonymity in the review process. Authors should prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

Author

JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Abstract. An abstract of not more than 200 words should identify the purpose of the article, provide an overview of the content, and suggest findings.

A list of references should be submitted on a separate page of the manuscript with the centered heading: *References*. The entries should be arranged alphabetically by surname of authors. The sample list of references below illustrates the format for bibliographic entries:

Dulay, H. & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 16* (1), 93-95.

Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

Reference citations in the text of the manuscript should include the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that are being quoted originally appeared, e.g., (Jones, 1982, pp. 235-238). All works cited in the manuscripts must appear in the list of references, and conversely, all works included in the list of references must be cited in the manuscript.

Notes should be used for substantive information only, and they should be numbered serially throughout the manuscript. They all should be listed on a separate page entitled *Notes*.

Specifications for Floppy Disks

Where feasible, manuscripts are preferred on Windows-based software. Manuscripts produced on one of the DOS-based or Macintosh systems should be formatted as an ASQII file at double density, if possible.

When mailing a floppy disk, please enclose the following:

1. Word processing software used: _____
2. Disk is formatted as: double__ high density__
3. Remarks:_____

Review Process

Manuscripts will be acknowledged by the editor upon receipt. Following preliminary editorial review, articles will be sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue.

• U.S. GOVERNMENT PRINTING OFFICE: 1997-585-380/79145

1142

1143

PIN: 074 768 - 000

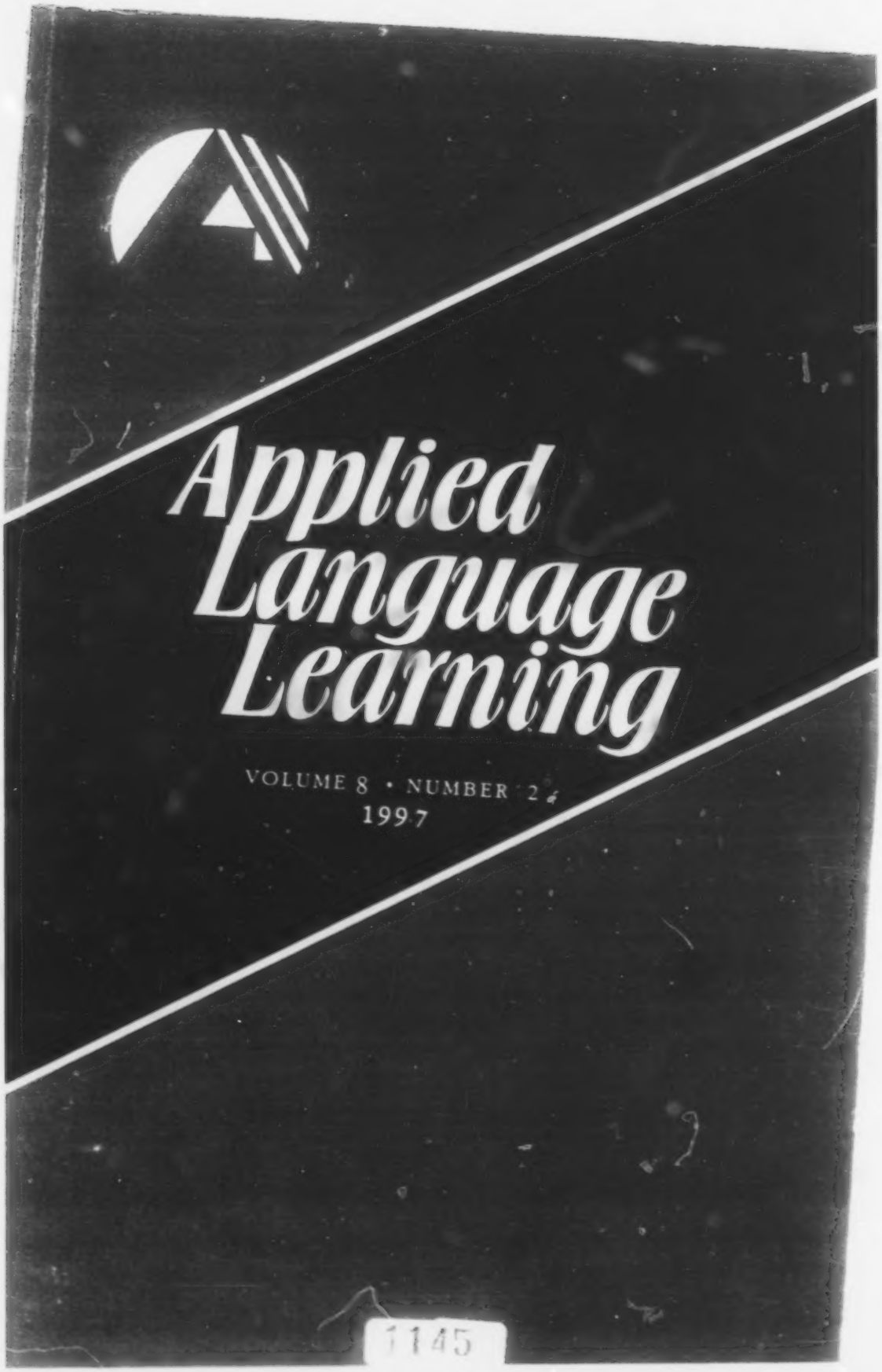
DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

BULK RATE
U.S. POSTAGE
PAID
MONTEREY, CA
PERMIT NO. 42

Center for Applied Linguistics
1118 22nd Street, NW
Washington, DC 20037
USA

PB-65-97-1
United States Army
PIN: 074 768-000
Approved for public release.
Distribution is unlimited.

1144


Applied Language Learning
VOLUME 8 • NUMBER 2
1997
1145

Applied Language Learning

Professional Bulletin 65-97-2

1997 **Volume 8** **Number 2**

Lidia Woytak
Editor

Joseph Morgan
Copyeditor

Defense Language Institute Foreign Language Center

Presidio of Monterey, CA 93944-5006

1146

Applied Language Learning

PB 65-97-2

The mission of Professional Bulletin 65-97, *Applied Language Learning (US* ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semiannually by the Defense Language Institute Foreign Language Center and Presidio of Monterey, presents professional information. The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

DENNIS J. REIMER
General, United States Army
Chief of Staff

Official:

JOEL B. HUDSON
Administrative Assistant to the Secretary of the Army
04239

DANIEL D. DEVLIN
Colonel, United States Army
Commander
DLIFLC and POM

EUGENE F. BEAUVAIS
Colonel, United States Air Force
Assistant Commandant
DLIFLC

1147

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the authors.

Availability

To access *Applied Language Learning* on the Internet type

http://lingnet.army.mil

Additionally, you may obtain the journal on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster

Send change-of-address information to

Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006

United Parcel Customers

Identify location of *Applied Language Learning*
Tin Barn
Bldg. 518, Room 6

Readers

Contact Editor, Dr. Woytak, by:

E-mail: woytakl@pom-emh1.army.mil
Telephone: (408) 242-5638
DSN: 878-5638
Fax: (408) 242-5850

Wordprocessing and graphics

Carol Liles and ***PV2 Lamar Davies***

Cover design

Barney Inada and ***PV2 Lamar Davies***

1148
U61

Reviewers for *Applied Language Learning*

The individuals listed below served as reviewers of manuscripts submitted to *Applied Language Learning* in 1997. We express our gratitude for expert service to:

Kathleen M. Bailey
Monterey Institute of International Studies
Martha S. Bean
San José State University
Christine M. Campbell
Defense Language Institute Foreign Language Center
John B. Carroll
University of North Carolina
Marianne Celce-Murcia
University of California, Los Angeles
Dan Douglas
Iowa State University
Rod Ellis
Temple University
Roderic A. Gale
Defense Language Institute Foreign Language Center
Robert C. Gardner
University of Western Ontario
John S. Hedgcock
Monterey Institute of International Studies
Evelyn Hatch
University of California, Los Angeles
Eli Hinkel
Seattle University
Gordon Jackson
Defense Language Institute Foreign Language Center
Ruth Johnson
Southern Illinois University
Gisèle Kapuscinski
Monterey Institute of International Studies
Dorry M. Kenyon
Center for Applied Linguistics
Stephen D. Krashen
University of Southern California
B. Kumaravadivelu
San José State University
James P. Lantolf
Cornell University
James F. Lee
University of Indiana
Virginia LoCastro
International Christian University
Steven Loughrin-Sacco
Boise State University
Barbara Matthies
Iowa State University
Jeffrey McQuillan
University of Southern California
Paul Nation
Victoria University of Wellington
Martha Nyikos
Indiana University
Rebecca Oxford
University of Alabama
Richard Schmidt
University of Hawaii-Manoa

Ilsoon Shin
Defense Language Institute Foreign Language Center
David J. Shook
Georgia Institute of Technology
Fredricka L. Stroller
Northern Arizona University
John Thain
Defense Language Institute Foreign Language Center
Roberta J. Vann
Iowa State University
Swathi Vanniarajan
San José State University
Leo Van Lier
Monterey Institute of International Studies
Keith Willsey
Defense Language Institute Foreign Language Center
Dolly J. Young
University of Tennessee

1150

Applied Language Learning

PB 65-97-2

1997 Volume 8 Number 2

Guest Editorial

147 Army Language Needs for the New Century
General Dennis J. Reimer

Articles

151 The Effects of Input Enhancement and Text Length on Adult L2 Readers' Comprehension and Intake in Second Language Acquisition
Ronald P. Leow

183 An Interactive Model of Vocabulary Acquisition
Swathi Vanniarajan

217 Pronunciation Instruction for "Fossilized" Learners: Can It Help?
Tracey M. Derwing, Murray J. Munro, and Grace Wiebe

237 What Researchers Say and Practitioners Do: Perspectives on Conscious Grammar Instruction in the ESL Classroom
Miriam Eisenstein Ebsworth and C. William Schweers

261 Proficiency Guidelines and Language Curriculum: Making ACTFL Proficiency Guidelines Effective in Furthering Japanese Language Proficiency
Yuki Johnson

Research Report

285 Voices from the Classroom: Intermediate-Level French Students Favor Extensive Reading over Grammar and Give Their Reasons
Beatrice C. Dupuy

151

News and Views

295 Linguist in Action: Interview with Colonel Daniel D. Devlin

Reviews

317 Strange (Ed.): Penguin Readers.....*Paul Nation*
323 Eggins and Slade: Analysing Casual Conversation.....*Martha S. Bean*
327 Taylor and Haas: German: A Self-Teaching Guide.....*Brigitte Olsen*
329 Vuong and Moore: Colloquial Vietnamese.....*Nguyen Dinh-Hoa*

General Information

333 Calendar of Events
339 Information for Contributors

1152

Guest Editorial

from the US Army Chief of Staff

Army Language Needs for the New Century

The Army's requirements for linguists in the 21st century reflect the dynamic, changing world in which we live. For the United States Army the new century began in 1989 with the dramatic end of the Cold War. Before that time we lived in a dangerous, but a relatively predictable world. The Soviet threat defined our doctrine, force structure, and our language training needs. Today we live in a world that is equally dangerous, but far less predictable. We have a new national security strategy that requires the Army to be globally engaged in order to shape the environment. Since the end of the Cold War the United States Army has participated in 28 major deployments around the globe. In 1997, in addition to our forward deployed troops, the Army had on average 31,000 soldiers posted to over 70 countries. As our requirements have grown, they have also diversified. We have seen a need for new languages and increased proficiency with less preparation and training time. In addition to language support for traditional military tasks, the Army also has expanding requirements for civil/military operations, host-nation coordination for logistical support, multinational military liaison, medical interpreters, legal translators, psychological operations, and military police operations. Faced with the vast challenges and opportunities of today's international environment the Army can no longer predict with certainty where and when the Army will need linguists. This is a Total Army—Active, Guard, and Reserve—challenge. To meet the demands of the next century, the Army is expanding and broadening our language training and support programs to ensure we are well prepared for the tasks ahead.

Since 1941 the Defense Language Institute Foreign Language Center (DLIFLC) and its predecessors have provided the foundation of the Army's language training program. We expect that the DLIFLC will continue to serve as the core of this effort into the next century. The DLIFLC consistently produces well-trained linguists for the military Services. At the rate of nearly Army 2,000 linguists per year, they have a proven track record that cannot be matched by any other institution in the US. Currently, they are working to further improve

academic performance, improving listening and reading abilities, cutting academic attrition and increasing the percentage of graduates achieving proficiency in all 21 languages taught at DLIFLC in Monterey, California.

We expect that the DLIFLC will continue to serve as our primary tool to build up the Army's inventory of linguists. In order to respond quickly to contingencies around the world, however, we must augment this process by tailored training delivered on demand, expanded sources of heritage linguists (those who enter the Army with already developed language skills), and the intelligent use of technology to improve training and proficiency of our limited inventory of linguists in the active force, Army Reserve, and the National Guard.

Part of the answer to this challenge will require new training methodologies. In particular, we'll require package sets of options to quickly establish flexible training programs. These training menus should facilitate conversion training to turn a linguist in one language into a capable linguist in a related language; augmentation training to bring a heritage linguist up to standard; refresher training to restore proficiency; maintenance and sustainment training; advanced training to increase proficiency, and cross-training to take a linguist from one language to a totally different language. The key to all these choices is rapid response.

We must take advantage of already existing resources. The United States had over 31 million people in 1990 over the age of five who spoke a foreign language, over 13% of our population. We need these heritage linguists. Many of these linguists could operate under conditions similar to our Individual Mobilization Augmentees (IMAs). Discussions are underway to increase the recruitment of heritage linguists into professional linguist positions and into other positions where their language skills can be used when needed. Earlier efforts to recruit heritage linguists already in the Army into professional linguist positions within Military Intelligence were successful. The DLIFLC can play a major role in providing augmentation training to these soldiers. Heritage speakers can also be used to supply civilian contract linguists. In situations where either the military inventory is inadequate or the job inappropriate for the military, contracted civilians from within the US can help fill the requirement.

Technological tools can and must aid in the training process. The use of appropriate tools to augment the natively fluent instructor can be a very innovative part of DLIFLC training. They have had an active program seeking new applications for technology since 1989 and have

1-154

a number of current computer based training approaches integrated into the training program. The DLIFLC has been a pioneer of distance learning. Two-way foreign language maintenance and sustainment training is conducted between the DLIFLC studios and troop units worldwide. They are currently exploring the applicability of desktop video to provide training to groups of three or fewer students.

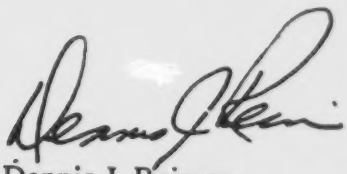
In addition to improving training methods we think new technology can also support operational needs. Simple word-for-word translators are already in use to sift large numbers of documents for analysis. A non-linguist can determine what needs to be translated. This frees the linguist for the more useful, and more difficult, job of translating. The Army has developed a system that combines a scanner, optical character recognition, and machine translation in a portable package. Another system, requiring only a CD-ROM and a laptop computer, allows a non-linguist to ask *yes*-or-*no* questions in a number of foreign languages to perform basic medical diagnosis. This system, combined with a wearable computer and English language voice recognition software, enables the operator to perform almost any sort of simple, *yes*-or-*no* interrogation. These kinds of rapidly adjustable, responsive systems will greatly enhance our ability to support contingency operations and unforeseen requirements.

We also anticipate leveraging the capabilities of civilian industry to enhance our operational capabilities. For example, AT&T's Language Line uses telephones and a computerized control center to link translators instantaneously with customers worldwide. Capabilities like the AT&T system have great potential for satisfying our needs and we are actively pursuing partnerships with industry.

In the years to come the Army will have unprecedented requirements for skilled linguists. These linguists will help coordinate with coalition forces, respond to humanitarian crises, deal with potential threats and shape the international environment through Army's many programs that foster cooperation and understanding. By combining a proven education system, innovative recruiting and personnel policies, new technology, and a strong relationship with industry; the Army is building a dynamic, responsive language support system for the next century. With the support of the DLIFLC, our partners in the academic community and industry, and commitment and dedication of American soldiers and our Department of the Army civilian employees, we will be well prepared to face any national security challenges that the future may hold.

1155

We will leverage technology in every way we can but at the end of the day it comes down to having an adequate number of Total Army soldiers and civilians trained for this important task. They are our credentials and DLIFLC will play a vital role in ensuring they are capable of accomplishing their vital mission in this new and unpredictable world.



Dennis J. Reimer
General, United States Army
Chief of Staff

1156

Articles

Applied Language Learning
1997, Vol. 8, No. 2, pp. 151-182

The Effects of Input Enhancement and Text Length on Adult L2 Readers' Comprehension and Intake in Second Language Acquisition

Ronald P. Leow
Georgetown University

This study investigates the effects of written input enhancement and text length on adult first-year second-language (L2) comprehension of text content and intake of targeted linguistic forms contained in the input.

Participants were 84 college-level second-semester students of Spanish who were exposed to one of four conditions: a long, non-enhanced text; a long, enhanced text; a short, non-enhanced text; and a short, enhanced text. To measure learners' comprehension and intake, a two-way analysis of variance (ANOVA) and a repeated-measures ANOVA were performed on the raw scores obtained on a comprehension task and a multiple-choice recognition assessment task respectively.

Results revealed a significant main effect for text length on readers' comprehension but not on intake, no significant main effect for input enhancement on either comprehension or intake, and no significant interaction. The findings provide empirical evidence that supports the validity and efficacy of exposing first-year L2 learners to short authentic reading materials to facilitate reading comprehension but not intake of linguistic items contained in the input. The findings also raise questions as to the relative effectiveness of external manipulation of written input to facilitate these learners' comprehension and further processing of linguistic forms in the input.

© *1997, Ronald P. Leow*

1157

The explosion of research on the second/foreign language reading process in the classroom reveals an increasing interest in investigating this process from a cognitive perspective (e.g., Anderson, 1991; Barnett, 1988; Block, 1986; Carrell, 1989; Hosenfeld, 1984; Hudson, 1982; Koda, 1988; O'Malley & Chamot, 1990; Oxford, 1992; Oxford & Ehrman, 1992; Steffensen & Joag-Dev, 1984). While these studies have focused mainly on identifying and classifying strategies L2 readers employ for comprehension purposes while interacting with the information in the text, only recently have there been a few published studies that have attempted to address the reading process under an attentional input processing framework (e.g., Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; Leow, 1993, 1995; Shook, 1994). The domain chosen for investigation has been the effects of textual manipulation (or more specifically, simplification and input enhancement) on readers' intake and production of linguistic forms contained in the input. However, the findings of these studies, due to limitations and methodological problems, need to be taken with some caution and underscore the need to further address the effects of written textual manipulation on second language acquisition (SLA). This study attempts to address the effects of input enhancement and text length on second semester L2 readers' comprehension of the text content and their intake of grammatical information contained in written input. Following Leow (1993), intake is defined in this study as "an intermediate process between the exposure to input and actual language acquisition ... that part of input that has been attended to by second language learners while processing the input" (p. 334).

Previous Research

Several theoretical approaches to the role of attention to formal features in the input in SLA have been postulated; for example, Long's (1991) "focus on form", McLaughlin's (1987) metaphor of limited-capacity processor of information, Schmidt's (1990, 1993, 1994, 1995) "noticing" hypothesis, Tomlin and Villa's (1994) model of input processing, and Robinson's (1995) model of the relationship between attention and memory. Underlying all these approaches is the general acceptance that some form of attention is crucial in promoting further processing of grammatical information in adult learners' L2 development (see also Ellis, 1993; Sharwood Smith, 1981, 1991). Current psycholinguistic studies have generally supported this role for attention in SLA (e.g., Cadierno, 1995; Doughty, 1991; Fotos, 1993; Lightbown & Spada, 1990; Spada & Lightbown, 1993; Trahey & White,

1158

1993; VanPatten & Cadierno, 1993; VanPatten & Sanz, 1995; White, Spada, Lightbown, & Ranta, 1991). In the written mode, one technique employed to promote readers' attention to linguistic items in the input has been textual manipulation of the input to which L2 learners are exposed. The two types of textual manipulation found in current literature are simplification (Leow, 1993, 1995) and input enhancement (Jourdenais *et al.*, 1995; Shook, 1990, 1994).[1]

Working under the metaphor of the L2 learner as being a limited-capacity processor of information (that is, adult L2 learners' amount of attention paid to L2 input depends heavily on the amount of cognitive effort required to process the input), Leow (1993, 1995) attempted to address the impact of simplification on adult learners' intake of grammatical forms contained in the input in both the written and the aural modes. Among his hypotheses was the following: A more comprehensible, simplified text should reduce the processing demands required to process for meaning, which in turn should facilitate learners' intake of the linguistic items under investigation. The texts were two modified versions of two authentic articles. That is, secondary information was omitted and linguistic changes were made to include more target items in the text. These versions were subsequently simplified to arrive at a statistically significant difference in comprehensibility between the original modified versions and the simplified versions. There were eight instances of the present perfect form in two of the four versions and eight instances of the present subjunctive form in the remaining two versions. The lengths of the four texts were 229, 237, 222, and 233 words. Participants were 137 (in the 1993 study) and 213 (in the 1995 study) college-level students enrolled in the first and fourth semester levels of study. They were exposed to written and aural input in one of the four conditions: a simplified or unsimplified text with the present perfect form, or a simplified or unsimplified text with the subjunctive form. Using repeated-measures multivariate analysis of variance (MANOVA) to analyze the data obtained on a multiple-choice recognition assessment task, I found that reducing learners' cognitive demands to processing for meaning through simplification did not produce any significant increase in their intake of the linguistic items contained in the input in either mode. In other words, I did not find any evidence that textual manipulation, in this case, simplification, facilitated learners' reallocation of their attention to form when their processing demands to attend to meaning were reduced.

Under the same theoretical framework of attention, Shook (1994) investigated the effects of type of saliency or input enhancement of grammatical information on learners' intake of two linguistic items,

1159

the present perfect and the relative pronouns (*que* and *quien*) in Spanish. One hundred and twenty-five second- and fourth-semester college-level students participated in this study. Like my studies, the texts used for both the present perfect and relative pronoun forms were two modified versions of two authentic articles. The lengths of the two texts were 185 and 217 words respectively and contained six present perfect tense forms and six relative pronoun forms. To promote readers' attention to the targeted forms in the input, Shook employed typographical cues (uppercasing and bolding). Participants were divided into three groups: The first was exposed to the text alone, the second was exposed to the grammatical items upper-cased and bolded with no instructions to pay attention to the targeted forms, and the third received a similar text like the second group together with the request to deduce a grammatical rule for the highlighted items. After exposure, participants completed a multiple-choice recognition task and a written fill-in-the-blank production task. Using repeated-measures MANOVA to analyze his data, Shook found significant effects for input enhancement of grammatical items on learners' intake and production of grammatical information contained in written input but only when instructed to pay attention to the targeted forms. However, in a relatively similar study, Shook (1990) found no significant difference in the three groups' performances on a free written recall task designed to measure their comprehension of the text content.

Jourdenais *et al.*, like Shook (1994), also employed typographic manipulation and typographic cues by bolding, shading, and underlining the forms in the written input. The text chosen for exposure was a merged narration of 204 words written originally by three native speakers of Spanish that included 18 preterit forms that were underlined and shadowed and 10 imperfect forms that were underlined and bolded. Both forms were also printed in a different and larger font to make these forms more salient to the readers. Participants were 10 college-level students enrolled in a second-semester class, half of whom were assigned to the enhancement group that received the enhanced text while the other half comprised the comparison group that received the non-enhanced version of the text. Participants' equivalent proficiency of the preterit and imperfect forms before exposure was measured by their performances on a midterm exam that included a fill-in-the-blank paragraph that elicited either tense based on the comprehension of the text. After exposure participants performed a written task in which they wrote a narration based on a series of pictures that described a robbery during Christmas. Participants were also requested to think aloud while performing this written task. Wilcoxon rank sums test per-

1160

formed on the data revealed that the enhancement group produced in their protocols more episodes related to selection and conjugation of the preterit and imperfect verbs and more target features on the written task when compared to the comparison group. The authors conclude that textual enhancement promotes noticing of target L2 forms and has an effect on subsequent output.

Summary and Limitations of Previous Research

While results from the above-cited studies appear to suggest that there may be some positive effects due to input enhancement, these findings need to be viewed with some caution due to several limitations. First of all, it is somewhat surprising that Jourdenais *et al.* (1995) and Shook (1994) did not measure or report the effects of input enhancement on learners' comprehension. Addressing learners' comprehension has both theoretical and practical implications. Theoretically, it is of interest to investigate whether the findings would support or refute the hypothesis that written input enhancement of L2 input facilitates both comprehension and intake in SLA. Pedagogically, the findings would address the practical aspect of enhancing written input in the classroom to promote comprehension and/or further processing of targeted forms.

Secondly, the number of targeted items also appears to be a factor to be considered. In Shook's study, there were six instances of each targeted form in the input. In Jourdenais *et al.*'s study, while participants were exposed to an increased number of targeted forms, there was quite an imbalance between the numbers of instances of the preterit forms (18) when compared to the imperfect forms (10). In addition, the low number of participants (5) in Jourdenais *et al.*'s experimental cells cannot be generalized to the population while there may be some doubt as to the validity of the pretest employed to measure learners' language ability before exposure to the enhanced and non-enhanced texts.

While the above-cited studies have shed some light on the effects of written input enhancement on readers' intake and output, the findings are not altogether conclusive and this line of research needs to be pursued. For example, the effects of text length also need to be addressed for its potential impact on readers' comprehension of text content and intake of linguistic forms in the input. As can be seen, the choice of written texts in these studies appears to have been motivated mainly by their type (expository) and content (topic familiarity and grammatical items)[2] while text length appears to have been consciously

1161

considered only for comparative purposes. From a psycholinguistic perspective, does the length of the text learners receive have an impact on what they comprehend and take in from the input? Can it be argued that the shorter the length of the text, the less difficult the task becomes resulting in less cognitive effort and attention when compared to a text of substantially greater length? Does this potential reduction of cognitive constraints then facilitate comprehension and intake, especially during the early stages of SLA? On the other hand, does exposure to substantially more input provide additional cues for learners to take in more content and linguistic information from the input? The issue of the role of text length on learners' comprehension and intake certainly warrants further investigation to shed additional insights into the L2 reading process.

Given our increasing understanding of the cognitive processes L2 readers employ and that reading materials in foreign language textbooks provide a substantial source of input to language learners, addressing written input enhancement and text length and their effects on comprehension of text content and intake of grammatical information is clearly warranted. The findings should provide (1) a clearer picture of what readers attend to while reading authentic L2 materials, and (2) empirical evidence to support or refute the current use of short written foreign language reading passages for pedagogical purposes and the validity and effectiveness of this practice in introductory foreign language textbooks.

Research Questions

The purpose of this study is to address the following specific research questions:

1. What are the effects of input enhancement on second-semester L2 readers' comprehension of the content information they receive in written authentic texts?
2. What are the effects of input enhancement on second-semester L2 readers' intake of new grammatical items contained in written authentic texts?
3. What are the effects of text length on second-semester L2 readers' comprehension of the content information they receive in written authentic texts?
4. What are the effects of text length on second-semester L2 readers' intake of new grammatical items contained in written authentic texts?

1162

Methodology

Participants

Participants were 84 native speakers of English enrolled in the second semester of Spanish study. These participants were not formally introduced to the targeted linguistic form under investigation and had received approximately 55 hours of formal exposure to Spanish prior to the experiment. Most of them had minimal prior knowledge of the targeted forms (see Analysis below). However, because reading was actively promoted in the curriculum, they were exposed to reading authentic texts of varying lengths in the classroom.

Targeted Linguistic Form

I chose the impersonal imperative Spanish forms for this study. The rationale for the choice of the imperative was two-fold: (1) while these forms were not formally presented to the participants before participating in the study, contextual guessing of these forms was highly possible due to the text type in which the imperative forms were found (see Materials below), and (2) comprehension of the text content as measured on the comprehension task (see Post-Exposure Assessment Tasks below) was dependent upon attending to these forms.

Materials

To address the effects of input enhancement and text length, I prepared four texts. The original text that I chose for the study was an article from the section titled *Medicina preventiva* ("Preventive Medicine") of a weekly Spanish magazine *mía* (1995, vol. *468*, p. 19) that provided several useful tips for leading a healthy life. I chose this text because it lent itself to the natural occurrence of imperatives while discussing a topic not unknown to the students. The original unmodified text (Text 1) contained 631 words in which there were 24 instances of naturally occurring imperative forms. The title was *Consejos útiles para llevar una vida saludable* ("Useful advice to lead a healthy life"). In this study, the group exposed to this unadulterated text was used as the control group (see Appendix A1). In addition to the original text, three versions of this text were prepared. The first version (Text 2) comprised the original text with one difference: 15 of the 24 instances of the imperative, found in the middle of the passage, were bolded and underlined (see Appendix A2). The second (Text 3) and the third (Text 4) versions were substantially shortened versions of Text 1 and Text 2

1163

respectively, comprising only the section of these texts that contained the 15 highlighted forms found in Text 2. In other words, Text 3 contained no textual enhancement while Text 4 did (see Appendix A3 and Appendix A4). The shortened versions (Texts 3 and 4) contained 384 words, just over half of Text 1 and Text 2.

Post-Exposure Assessment Tasks

To measure participants' comprehension, a written comprehension task was prepared. Participants were requested to answer 10 questions written in English (their native language) based on the reading passage to which they were just exposed. To maintain consistency in all participants' performances, the questions addressed only the information on the shortened versions. These 10 questions, requiring 20 pieces of information, were designed to address the tips, advice, or recommendations provided mainly through the use of the imperatives in the input. Because production of the foreign language was not an issue in this study, they also were required to answer in English. In order to control for passage-dependent answers, the comprehension task alone was then administered to 25 participants at the same level of language proficiency. Results indicated that out of the 20 answers on the task, 10 were relatively passage dependent, that is, participants needed to read the passage in order to provide the correct answer 97 percent of the time.[3] (see Appendix B1)

To measure participants' intake of the targeted forms in the input, a multiple-choice recognition assessment task was prepared (cf. Leow, [1993, 337-338] for the rationale for using such a task to measure intake). There were 15 items on this task that provided tips, pieces of advice, or recommendations for a healthy life taken directly from the text. Once again, these items addressed only the targeted forms contained in the shortened versions to maintain consistency in participants' performances. Participants were required to select out of four possibilities the one form that would grammatically provide the tip, piece of advice, or recommendation. As in Leow (1993, 1995), in order to avoid the extraneous variable of other knowledge sources influencing participants' immediate recognition of the targeted forms, participants were informed to complete the sentences in the order they appeared and not to skip a number and then return to it. This task comprised both the pretest and posttest with one important difference: The targeted forms in the texts were presented in different contexts on the pretest.[4] (see Appendix B2)

1164

Testing Procedure

Two weeks before exposure to the input, participants completed the pretest designed to measure their grammatical recognition of formal imperative forms.

On the day of the experiment, participants randomly received a packet that contained an agreement form, instructions, one of the four texts, and the comprehension and recognition tasks in that order. They were provided in the instructions a translation of the title of the article to activate prior knowledge and then asked to read the text at their own pace. They also were informed that they had been randomly assigned one of four versions of their article. After completing the text, they were requested to turn the page and, without looking back at the text, perform the two tasks in the order presented.

Scoring Procedure

The comprehension task was scored one point or none for each correct passage-dependent answer for a total of 10. For the recognition task, one point or none was scored for each correct choice for a total of 15. Table 1 and Table 2 display the mean scores and standard deviations for each experimental cell on the comprehension task and the pretest and posttest recognition tasks respectively.

TABLE 1
Mean Scores and Standard Deviations for Each Experimental Cell on the Comprehension Task

	Text	
Input enhancement	Short	Long
Non-enhanced	3.76	2.35
SD	2.24	1.50
	$n = 25$	$n = 20$
Enhanced	3.65	3.37
SD	1.98	1.83
	$n = 20$	$n = 19$

1165

TABLE 2
Mean Scores and Standard Deviations for Each Experimental Cell on the Pretest and Posttest Recognition Task

	Text			
	Short		Long	
Input enhancement	Pretest	Posttest	Pretest	Posttest
Non-enhanced	1.88	4.40	2.30	4.50
SD	1.56	4.15	1.72	4.30
	$n = 25$		$n = 20$	
Enhanced	2.30	4.10	2.10	4.37
SD	1.75	3.73	1.85	2.41
	$n = 20$		$n = 19$	

Results

To measure the effects of input enhancement and text length on participants' comprehension of the text, the raw scores obtained on the comprehension task were submitted to a 2 x 2 ANOVA using a two between-subject, zero within-subject factorial design. Input enhancement (enhanced vs. non-enhanced) and text length (short vs. long) were the two independent variables. The dependent variable were mean scores obtained on the comprehension task. All analyses were run on the SuperANOVA statistical package.

The results of the ANOVA revealed a significant main effect for text length (F [1, 84] = 3.99, p = .0493), no significant main effect for input enhancement (F [1, 84] = 1.15, p = .2868), and no significant interaction between input enhancement and text length (F [1, 84] = 1.774, p = .1867). The results indicate that participants exposed to the short versions comprehended significantly more the content of the text when compared to participants exposed to the long versions (3.71 vs. 2.85 respectively; p = .049). However, the results also indicate that participants exposed to the enhanced texts did not demonstrate significantly better comprehension of the text content when compared to those exposed to the non-enhanced texts (3.51 vs. 3.13, respectively; p = .287). The results of the ANOVA are presented in Table 3.

1166

TABLE 3
ANOVA Summary Table: Comprehension Task

Source of variation	df	Sum of squares	Mean square	*F*	*p*
Input enhancement	1	4.284	4.284	1.150	.2868
Text length	1	14.854	14.854	3.987	.0493*
Input enhancement x Text length	1	6.610	6.610	1.774	.1867
Error	80	298.081	3.726		

$N = 84$
$^*p < .05$

A two-way ANOVA performed on the total mean scores (out of a total of 20) obtained on the comprehension task also revealed similar statistical results. That is, the only significant main effect found was for text length (F [1, 84] = 5.34, p = .0234). In other words, participants exposed to the short versions significantly outperformed those exposed to the long versions (9.58 vs. 7.94 for the short and long versions, respectively). On the other hand, there was no difference in performance between participants exposed to the enhanced and non-enhanced versions (8.67 vs. 8.98 for the enhanced and non-enhanced versions respectively; $p > .05$).

To measure the effects of input enhancement and text length on participants' intake of the targeted linguistic forms in the input, the raw scores obtained on the multiple-choice assessment task were submitted to a 2 x 2 x 2 repeated-measures ANOVA using a two between-subject, one within-subject factorial design. Input enhancement (enhanced vs. non-enhanced) and text length (short vs. long) were entered as the two between-subject factors while time (pretest vs. posttest) was entered as the within-subject factor.

The results revealed no significant main effects for either input enhancement (F [1, 84] = .014, p = .7386) or text length (F [1, 84] = .112, p = .7386), nor was there any significant interaction between input enhancement and text length (F [1, 84] = .063, p = .8019). The results indicate that participants exposed to the enhanced texts did not demonstrate significantly better intake of the targeted forms when compared to those exposed to the non-enhanced texts (3.22 vs. 3.26, respectively; p = .908). The results also indicate that participants exposed to the short versions did not demonstrate superior intake of the

1167

targeted forms when compared to participants exposed to the long versions (3.17 vs. 3.32 respectively; $p = .739$). In other words, the results indicate that neither input enhancement nor text length facilitated participants' intake of the imperative forms contained in the input. The results of the ANOVA are presented in Table 4.

TABLE 4
ANOVA Summary Table: Recognition Task

Source of variation	df	Sum of squares	Mean square	*F*	*p*
Input enhancement	1	.110	.110	.014	.9077
Text length	1	.915	.915	.112	.7386
Input enhancement x Text length	1	.517	.517	.063	.8019
Error	80	652.888	8.161		
Time	1	200.237	200.237	22.399	.0001*
Time x Input enhancement	1	1.120	1.120	.125	.7243
Time x Text length	1	.053	.053	.006	.9387
Time x Input enhancement x Text length	1	1.592	1.592	.178	.6742
Error	80	715.162	8.940		

N=84

The results of the ANOVA also revealed a significant main effect for time (F [1, 84] = 22.399, $p = .0001$). In other words, participants demonstrated an overall improvement in performance on their posttest when compared to their performance on the pretest (4.34 vs. 2.13 for the posttest and pretest, respectively).

1168

ANOVAs performed on the pretest and posttest data separately also revealed no significant main effect for either input enhancement or text length ($p > .05$ for both the pretests and posttests).

Discussion

The results found for research question one that addressed the effects of input enhancement on second semester L2 readers' comprehension revealed that input enhancement did not appear to have any significant impact on learners' comprehension of the text content when compared to learners who did not receive this kind of input enhancement. In other words, highlighting the imperative forms in the input through the use of typographic cues, that is, bolding and underlining, did not appear to be effective in focusing readers' attention to these targeted forms to comprehend the passage. Recall that most of the questions on the comprehension task were based on the information provided by the imperatives in the input. The lack of significant difference in comprehension between the enhanced and non-enhanced texts appears to corroborate Shook's (1990) study that found similar non-significant results on a free written recall task designed to measure comprehension.

The answer to research question two that addressed specifically the effects of input enhancement on readers' intake was also negative: Enhancing the targeted forms through typographic cues did not appear to promote further processing of these forms as measured on a recognition task administered immediately after exposure when compared to learners not exposed to these enhanced forms in the input. While these findings appear to contradict both Shook's (1994) and Jourdenais *et al*'s (1995) findings that input enhancement promotes attention to or noticing of targeted L2 forms in the input, the type of instruction, post-exposure tasks, and type of linguistic forms need to be considered. In Shook's study, learners were explicitly instructed to pay attention to the enhanced forms in the input and their attention was measured on a multiple-choice recognition task. A similar post-exposure task was used in the present study but learners were only instructed to read for comprehension. However, while both the present study's and Jourdenais *et al.*'s learners received generally the same instruction to read for comprehension, the post-exposure tasks employed in the two studies measured different performances, namely, intake and output, respectively.

With respect to type of linguistic forms, while the present study investigated learners' attention to Spanish imperative forms, Shook investigated the Spanish present perfect forms and Spanish relative

1169

pronouns and Jourdenais *et al.* the Spanish preterit and imperfect forms. It may be argued that the different types of linguistic forms could have played a role in the contradicting results found in these studies. For example, forms such as the imperatives and the subjunctive may not be as meaning-bearing as the preterit and imperfect forms.

On the other hand, the non-significant effect of input enhancement on readers' intake corroborates Leow (1993, 1995) and White (forthcoming) who found similar results in their studies. Leow simplified input to which learners were exposed but found that this type of textual manipulation of the input did not appear to facilitate learners' intake of the present perfect and subjunctive forms in either the written or aural mode. White's (forthcoming) textual input enhancement (using enlargement and a different combinations of italics, bolding, and underlining) also appeared to be ineffectual in drawing grade six ESL francophone readers' substantial attention to the targeted forms in the input. She reported that her participants' responses to a questionnaire on the typographical enhancement indicated that not many learners were certain about the purpose of the enhanced forms. In addition, they also reported that enhancement was not useful in figuring out the grammatical rule under investigation.

The findings of the present study, together with my previous studies (1993, 1995), Shook's (1990), and White's (forthcoming) appear to suggest that textual manipulation without explicit instruction to attend to the targeted forms does not facilitate L2 readers' comprehension or intake of textual or grammatical information. This non-significant effect of external manipulation in SLA may need to be considered in light of the difference between theory and pedagogy. Theoretically, the hypothesis or claim that linguistic forms need to be noticed or detected before they can be further processed is an internally driven perspective that is generally accepted in SLA literature (e.g., Long, 1991; Schmidt, 1990, 1993, 1994, 1995; Tomlin & Villa, 1994). Pedagogically, however, external manipulation of input in the written mode designed to enhance learners' attention to targeted linguistic forms in the input represents an externally-driven perspective and may be more problematic. For example, external manipulation may not have its desired effects due to methodological factors that need to be considered and tightly controlled, such as (1) the amount or level of attention learners pay to the enhanced forms; (2) the role of awareness in processing input into intake; (3) the amount of exposure to the enhanced forms; and (4) the potential variation of learners' overall L2 proficiency found in experimental cells.

First of all, the amount or level of attention learners pay to the

1170

enhanced forms in the input is generally uncontrolled and highly speculative, since it is methodologically difficult to account for the role of attention during interaction with the input. More robust research designs are needed to first establish that attention was indeed paid to the enhanced forms before any post-exposure claims can be made with regard to the effects of attention to enhanced forms in the input.[5]

Second, there is considerable controversy as to the role of awareness in the human attentional system while learning a foreign or second language (e.g., Hardcastle, 1993; Joordens & Merikle, 1992; McLaughlin, 1990; Schmidt, 1990, 1993, 1994, 1995; Tomlin & Villa, 1994). For example, Schmidt argues that consciousness in the sense of awareness of specific items in the input at the level of noticing (or what is also referred to as conscious attention) is crucial for the conversion of input into intake. Tomlin and Villa, on the other hand, suggest that the role of conscious attention may not be important for language learning. They propose that it is the attentional function of detection (which can be dissociated from awareness) that is crucial for further processing to take place. What role awareness plays in facilitating enhanced input into intake also needs to be empirically investigated.[6]

Third, exposing L2 learners to one exposure of enhanced targeted forms in written input without any explicit instruction to attend to these forms may not provide sufficient time to draw their attention to these targeted forms in any substantial way. I (forthcoming-b) empirically investigated the effects of amount of exposure (single vs. multiple) on first-semester learners' intake and written production of Spanish irregular third-person singular and plural forms of stem-changing *-ir* verbs. I found that learners exposed to multiple exposures performed significantly better than learners exposed to only one exposure, a performance that lasted over one semester.

Fourth, a review of the raw scores in the present study for each group reveals quite a range of performance in individual gain scores. For example, in the short, non-enhanced group, there was a range of -4 to + 14; in the short, enhanced group, -5 to +14; in the long, non-enhanced group, -6 to +14, and in the long, enhanced group, -3 to +7. This variation appears to support my (1993) contention that "external manipulation of the input may not only be haphazard but also inadequate to address what may appropriately facilitate learners' intake" (p. 342) and Sharwood Smith's (1991: 121) statement that externally induced saliency may not necessarily be registered by learners who receive this type of information.

The results found for research question three that addressed the

1171

effects of text length on readers' comprehension of text content suggest that text length plays an important role in facilitating second semester college-level students' comprehension of authentic written text. Learners exposed to the short versions of the text comprehended the text content significantly more than those exposed to the longer versions. These findings lend empirical support to the current perception of textbook writers and publishing companies that written short texts are important for facilitating first-year L2 readers' comprehension in reading the foreign language. In addition, the findings can also help language teachers make informed pedagogical decisions on their selection of reading materials at this first-year level.

On the other hand, the findings do not indicate that text length has a facilitating effect on readers' intake of grammatical forms contained in the input. Consequently, the answer to research question four is negative. Learners exposed to either the long or short versions did not demonstrate any statistical difference in their performances on the recognition task. It appears that additional exposure to the targeted forms — that is, nine more occurrences of the imperative forms in the study — did not appear to make any substantial contribution to readers' intake of the targeted forms in the input. One plausible explanation may be that there are two separate processes occurring: one for comprehension and another for linguistic intake. Further studies on this possibility are warranted in order to shed light on such potential processes while interacting with written input.

Future Research

Before any definitive statement can be made on the effects of input enhancement on learners' attention to enhanced targeted forms in written input, there is need for more robust research designs that take into account an adequate operationalization of the process of attention in order to truly address what learners attend to, focus on, or notice in the input. To shed more light on the effects of input enhancement on learners' attention to targeted forms contained in the input, future research also needs to address the following empirical questions, among many others:

1. What role does text type play in the effects of input enhancement? In other words, does the format of written input have an effect on whether learners attend to enhanced forms in the input?
2. What is the effect of complex versus simple enhanced forms or meaning-bearing versus non-meaning enhanced forms? In other words,

do learners attend to some enhanced forms but not to others?

3. What is the effect of amount of exposure? That is, should learners be exposed to the input with the enhanced forms more than once in an effort to increase their attention to these forms?

4. Does level of language ability play a role in learners' attention to enhanced input?

Conclusion

This study has provided empirical evidence that supports the validity and efficacy of exposing first-year L2 learners to short authentic reading materials to facilitate reading comprehension, but not intake, of Spanish imperative forms contained in the input. Similar empirical support was not found for the effects of input enhancement on either readers' comprehension or intake, lending support to my (1993) argument that it may be "the learner's existing language system that defines what in the input is taken in and not any external manipulation of the input" (p. 342). It was also suggested that externally manipulated written input may not have its desired effects due to the methodological difficulty of tightly controlling the amount or level of readers' attention and awareness while interacting with the input, the amount of exposure, and the level of L2 proficiency in the experimental cells. Consequently, until further studies explore these variables, findings of the effect of written input enhancement on readers' comprehension and intake need to be taken with some caution.

Notes

[1] Simplification, as in Leow (1993, 1995), is defined as modifications (e.g., lexical, phonological, syntactic etc.) made to L2 input in an effort to facilitate L2 learners' comprehension. Input enhancement is defined as attempts to draw L2 learners' attention to targeted forms in the input by highlighting or making salient these forms through the use of typographic manipulation and typographic cues (e.g., bolding, shading, underlining, uppercasing etc.).

[2] This choice is applicable to most studies investigating the effects of text characteristics in the L2 reading process (e.g., Davies, 1984; Davis, 1989; Taylor, 1983) that have focused on their potential effects on readers' comprehension.

1123

[3] The passage-dependent items were the following answers: family eating needs (Q1); climb up the stairs and climb down the stairs (Q3); don't expose yourself to the sun for too long (Q5); use a helmet while riding a motorcycle or bicycle, drive rested and drive without listening to music (Q7); get vaccinations, take care of sexual hygiene, and take precautions while traveling abroad (Q10).

[4] A sample of the pretest is as follows:

Pretest

Here is a list of tips or advice for a healthy life. For each sentence, circle the letter *which, in your opinion, grammatically provides the tip, piece of advice, or recommendation.*

1. (Find out)_____ de sus habilidades (abilities) físicas.

A. Entérate B. Entérase
C. Entérese D. Entérete

2. (Make) _____ una lista de sus necesidades esenciales para hacer ejercicios.

A. Hace B. Haz
C. Haces D. Haga

3. (Paste)_____ la lista a la puerta de su cuarto.

A. Pegue B. Pega
C. Pegues D. Pegas

[5] In a study that addressed the operationalization of the process of attention in SLA, I investigated from a task-based approach and at a morphological level, Tomlin and Villa's (1994) fine-grained analysis of attention in SLA (forthcoming-a). In a tightly controlled study that first established that attention was paid to the targeted forms, I investigated the performances of four groups of beginning learners of Spanish who completed one of four crossword puzzles designed to isolate the effects of alertness, orientation, and detection. Repeated-measures ANOVAs performed on the raw scores obtained on the pretest and three posttests of a recognition and written production task revealed significant main effects for type of attentional function, time, and significant interactions. Results lend strong empirical support to Tomlin

1174

and Villa's (1994) fine-grained analysis of attention while indicating short term effects of detection.

[6] I provide positive effects for the role of awareness on learners' intake (1997). In this study, I qualitatively and quantitatively investigated the think-aloud protocols produced by beginning adult L2 learners of Spanish completing a problem-solving task and their immediate performances on two post-exposure assessment tasks, a recognition and written production task. The qualitative and quantitative analyses of learners' performances suggest the following conclusions: (1) different levels of awareness lead to differences in processing, (2) more awareness appears to contribute to more recognition and accurate written production of noticed forms, and (3) the findings of this study provide empirical support for the facilitative effects of awareness in foreign language behavior.

References

Anderson, N. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal*, *75*, 460-472.

Barnett, M. (1988). Teaching reading strategies: How methodology affects language course articulation. *Foreign Language Annals*, *21*, 109-119.

Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, *20*, 463-494.

Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal*, *79*, 179-193.

Carrell, P. (1989). Metacognitive awareness and second language reading. *Modern Language Journal*, *73*, 121-134.

Davies, A. (1984). Simple, simplified and simplification: What is authentic? In J. Alderson & A. Urquhart (Eds.), *Reading in a foreign language* (pp. 181-195). New York: Longman.

Davis, J. (1989). Facilitating effects of marginal glosses on foreign language reading. *Modern Language Journal*, *73*, 41-48.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition*, *13*, 431-469.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, *27*, 91-113.

Fotos, S. (1993). Consciousness raising and noticing through focus on form: Grammar task development versus formal instruction. *Applied Linguistics*, *14*, 385-407.

175

Hardcastle, V. G. (1993). The naturalists vs. the sceptics: The debate over a scientific understanding of consciousness. *Journal of Mind and Behavior, 14*, 27-50.

Hosenfeld, C. (1984). Case studies of ninth grade readers. In J. Alderson & A. Urquhart (Eds.), *Reading in a foreign language* (pp. 231-249). New York: Longman.

Hudson, T. (1982). The effects of induced schemata on the 'short circuit' in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning, 32*, 1-31.

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing?: A think aloud protocol analysis. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Technical Report #9, pp. 183-216). Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Joordens, S., & Merikle, P. M. (1992). False recognition and perception without awareness. *Memory and Cognition, 20*, 151-159.

Koda, K. (1988). The cognitive process in second language reading: Transfer of L1 reading skills and strategies. *Second Language Research, 4*, 133-156.

Leow, R. P. (1993). To simplify or not to simplify: A look at intake. *Studies in Second Language Acquisition, 15*, 333-355.

Leow, R. P. (1995). Modality and intake in second language acquisition. *Studies in Second Language Acquisition, 17*, 79-89.

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning, 47*, 467-506.

Leow, R. P. (forthcoming-a). Toward operationalizing the process of attention ins second language acquisition: Evidence for Tomlin & Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics, 19.*

Leow, R. P. (forthcoming-b). The effects of amount and type of exposure on adult learners' L2 development in SLA. *Modern Language Journal, 82*(1).

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Philadelphia: John Benjamins.

McLaughlin, B. (1990). "Conscious" vs. "unconscious learning". *TESOL Quarterly, 24*, 617-634.

O'Malley, J. M., & Chamot, A. (1990). *Learning strategies in second-language acquisition*. Cambridge: Cambridge University Press.

1176

Oxford, R. (1992). Research on second language learning strategies. *Annual Review of Applied Linguistics*, *13*, 175-187.

Oxford, R., & Ehrman, M. (1992). Second language research on individual differences. *Annual Review of Applied Linguistics*, *13*, 188-205.

Robinson, P. (1995). Attention, memory, and the 'noticing' hypothesis. *Language Learning*, *45*, 283-331.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206-226.

Schmidt, R. W. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. In J. H. Hulstijn & R. W. Schmidt (Eds.), *AILA Review: Consciousness and second language learning: Conceptual, methodological and practical issues in language learning and teaching*, *11*, 11-26.

Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Technical Report #9, pp. 1-63). Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research*, *17*, 118-136.

Shook, D. J. (1990). The processing of grammatical structures via written input. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon, *Applied Language Learning*, *5*, 57-93.

Spada, N. M. & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, *15*, 205-224.

Steffensen, M., & Joag-Dev, C. (1984). Cultural knowledge and reading. In J. Alderson & A. Urquhart (Eds.), *Reading in a foreign language* (pp. 48-61). New York: Longman.

Taylor, K. (1983). Can college students summarize? *Journal of Reading*, *26*, 254-258.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*, 183-203.

1177 1187

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*, 181-203.

VanPatten, B., & Cadierno, T. (1993). Input processing and second language acquisition: A role for instruction. *Modern Language Journal, 77*, 45-57.

VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R. Rutkowski (Eds.), *Second language acquisition: Theory and pedagogy* (pp. 169-185). Hillsdale, NJ: Lawrence Erlbaum.

White, J. (forthcoming). Getting the learners' attention: A typographical input enhancement study. To appear in C. Doughty & J. Williams (Eds.), *Focus on form.* Cambridge: Cambridge University Press.

White, L., Spada, N. M., Lightbown, P. M., & Ranta, L. (1991). Input enhancement and second-language question formation. *Applied Linguistics, 4*, 416-432.

Appendix A1

Text 1: Long Non-Enhanced Text

Medicina preventiva

Consejos útiles para llevar una vida saludable*

Llevar una vida sana no consiste solamente en adoptar una serie de hábitos alimenticios e higiénicos que se consideren poco perjudiciales para nuestra salud. También pase por elegir el médico de cabecera ideal (aunque le cueste dar varias veces la vuelta a su ciudad en su busca), saber reaccionar ante situaciones de emergencia, tener claro cuándo debe acudir al especialista, cumplir a rajatabla los consejos de su médico, proteger su piel, sus ojos y su esqueleto de agresiones externas, aprender a tomarse cada acontecimiento con calma y filoso fía ... En resumidas cuentas, adopte una actitud positiva ante la vida, tome sus propias riendas y procure preservar al máximo el capital de salud con el que usted y los suyos cuentan.

Desde luego, tiene que comer bien. Acostúmbrese a pensar antes de abrir la nevera: cada día tome fruta y verduras, carne o pescado, dos o tres huevos por semana, leche o queso a diario ... Entérese muy bien de sus necesidades alimenticias y también de las de su familia (sobre todo si tienè hijos pequeños). Haga una lista y péguela en la puerta de

la nevera. Y aunque de vez en cuando se puede permitir un capricho, adquiera zanahorias o incluso algún fruto seco mejor que bollos o pan.

El ejercicio físico es también imprescindible. No tiene que convertirse en una gran atleta; simplemente suba y baje las escaleras de su casa y vaya por todas partes a pie y su corazón tendrá la agilidad que necesita.

Ni beber, ni fumar. Sabe perfectamente cuáles son los hábitos que le perjudican, pero otros no son tan obvios. Como trasnochar a menudo. Aunque de vez en cuando sea necesario trabajar hasta altas horas de la noche o salir con los amigos para echar una canita al aire, duerma sin interrupciones entre 7 y 8 horas para que su organismo pueda recuperarse. Tampoco se exponga al sol durante demasiado tiempo sin protección, tanto en la piel como en los ojos. Otros hábitos importantes tienen que ver con su seguridad: use casco si va en moto o bici y conduzca siempre descansado y, mejor, sin música.

Puede evitar muchas enfermedades. En ocasiones, existen en las familias antece dentes de enfermedades a las que usted puede tener propensión. Pregunte a sus mayores la historia clínica y comente con su médico de cabecera si usted o alguno de los suyos podrían desarrollar alguna alteración en el futuro y qué medios podría poner para evitarlo. En cualquier caso, un chequeo bianualanual a partir de los 35-40 añosle será de gran ayuda para detectar posibles anomalias antes de que representen un serio peligro. ¡Y desde luego evite las situaciones de riesgo! Siga a rajatabla los calendarios de vacunación; tome medidas de higiene sexual tanto si no tiene pareja estable como si la tiene, y evite viajar a otras zonas del mundo sin tomar las precauciones básicas: vacunación si es necesaria, beber agua mineral, no comer nada en la calle.

Busque un buen médico y sea un buen paciente. No le serviría de nada que le atienda un premio Nobel si después es incapaz de seguir sus recomendaciones. Por ello, busque el médico generalista que le inspire confianza, le explique bien lo que le pasa, esté dispuesto a atender una situación de emergencia y le recomiende el especialista adecuado cuando la situación lo requiera. Un buen médico de cabecera es una institución fundamental para la salud y el bienestar de toda la familia. Y siga todas sus indicaciones, sin diagnosticar ni recetar tratamientos por su cuenta y riesgo.

Ríase de todo (siempre que pueda). El buen humor y el optimismo son excelentes preventivos tanto físicos como mentales.

C.G

Próxima semana
Todo lo que debe saber del cerebro

**MIA, 468.* (1995). Reprinted with Permission.

Appendix A2
Text 2: Long Enhanced Text

Medicina preventiva

Consejos útiles para llevar una vida saludable

Llevar una vida sana no consiste solamente en adoptar una serie de hábitos alimenticios e higiénicos que se consideren poco perjudiciales para nuestra salud. También pase por elegir el médico de cabecera ideal (aunque le cueste dar varias veces la vuelta a su ciudad en su busca), saber reaccionar ante situaciones de emergencia, tener claro cuándo debe acudir al especialista, cumplir a rajatabla los consejos de su médico, proteger su piel, sus ojos y su esqueleto de agresiones externas, aprender a tomarse cada acontecimiento con calma y filosofía ... En resumidas cuentas, adopte una actitud positiva ante la vida, tome sus propias riendas y procure preservar al máximo el capital de salud con el que usted y los suyos cuentan.

Desde luego, tiene que comer bien. **Acostúmbrese** a pensar antes de abrir la nevera: cada día **tome** fruta y verduras, carne o pescado, dos o tres huevos por semana, leche o queso a diario ... **Entérese** muy bien de sus necesidades alimenticias y también de las de su familia (sobre todo si tiene hijos pequeños). **Haga** una lista y **pegue** la lista en la puerta de la nevera. Y aunque de vez en cuando se puede permitir un capricho, **adquiera** zanahorias o incluso algún fruto seco mejor que bollos o pan.

El ejercicio físico es también imprescindible. No tiene que convertirse en una gran atleta; simplemente **suba** y **baje** las escaleras de su casa y vaya por todas partes a pie y su corazón tendrá la agilidad que necesita.

Ni beber, ni fumar. Sabe perfectamente cuáles son los hábitos que le perjudican, pero otros no son tan obvios. Como trasnochar a menudo. Aunque de vez en cuando sea necesario trabajar hasta altas horas de la noche o salir con los amigos para echar una canita al aire, **duerma** sin interrupciones entre 7 y 8 horas para que su organismo pueda recuperarse. Tampoco se exponga al sol durante demasiado tiempo sin protección, tanto en la piel como en los ojos.

1180

Otros hábitos importantes tienen que ver con su seguridad: **use** casco si va en moto o bici y **conduzca** siempre descansado y, mejor, sin música.

Puede evitar muchas enfermedades. En ocasiones, existen en las familias antecedentes de enfermedades a las que usted puede tener propensión. **Pregunte** a sus mayores la historia clínica y **comente** con su médico de cabecera si usted o alguno de los suyos podrían desarrollar alguna alteración en el futuro y qué medios podría poner para evitarlo. En cualquier caso, **obtenga** un chequeo bianual - anual a partir de los 35-40 años - le será de gran ayuda para detectar posibles anomalias antes de que representen un serio peligro. ¡Y desde luego **evite** las situaciones de riesgo! Siga a rajatabla los calendarios de vacunación; tome medidas de higiene sexual tanto si no tiene pareja estable como si la tiene, y evite viajar a otras zonas del mundo sin tomar las precauciones básicas: vacunación si es necesaria, beber agua mineral, no comer nada en la calle.

Busque un buen médico y sea un buen paciente. No le serviría de nada que le atienda un premio Nobel si después es incapaz de seguir sus recomendaciones. Por ello, busque el médico generalista que le inspire confianza, le explique bien lo que le pasa, esté dispuesto a atender una situación de emergencia y le recomiende el especialista adecuado cuando la situación lo requiera. Un buen médico de cabecera es una institución fundamental para la salud y el bienestar de toda la familia. Y siga todas sus indicaciones, sin diagnosticar ni recetar tratamientos por su cuenta y riesgo.

Ríase de todo (siempre que pueda). El buen humor y el optimismo son excelentes preventivos tanto físicos como mentales.

C.G

Próxima semana
Todo lo que debe saber del cerebro

Appendix A3

Text 3: Short Non-Enhanced Text

_____

Medicina preventiva _____

Consejos útiles para llevar una vida saludable

Desde luego, tiene que comer bien. Acostúmbrese a pensar antes de abrir la nevera: cada día tome fruta y verduras, carne o pescado, dos o tres huevos por semana, leche o queso a diario ... Entérese muy bien de sus necesidades alimenticias y

1181

también de las de su familia (sobre todo si tiene hijos pequeños). Haga una lista y péguela en la puerta de la nevera. Y aunque de vez en cuando se puede permitir un capricho, adquiera zanahorias o incluso algún fruto seco mejor que bollos o pan.

El ejercicio físico es también imprescindible. No tiene que convertirse en una gran atleta; simplemente suba y baje las escaleras de su casa y vaya por todas partes a pie y su corazón tendrá la agilidad que necesita.

Ni beber, ni fumar. Sabe perfectamente cuáles son los hábitos que le perjudican, pero otros no son tan obvios. Como trasnochar a menudo. Aunque de vez en cuando sea necesario trabajar hasta altas horas de la noche o salir con los amigos para echar una canita al aire, duerma sin interrupciones entre 7 y 8 horas para que su organismo pueda recuperarse. Tampoco se exponga al sol durante demasiado tiempo sin protección, tanto en la piel como en los ojos. Otros hábitos importantes tienen que ver con su seguridad: use casco si va en moto o bici y conduzca siempre descansado y, mejor, sin música.

Puede evitar muchas enfermedades. En ocasiones, existen en las familias antecedentes de enfermedades a las que usted puede tener propensión. Pregunte a sus mayores la historia clínica y comente con su médico de cabecera si usted o alguno de los suyos podrían desarrollar alguna alteración en el futuro y qué medios podría poner para evitarlo. En cualquier caso, un chequeo bianual - anual a partir de los 35-40 años - le será de gran ayuda para detectar posibles anomalias antes de que representen un serio peligro. ¡Y desde luego evite las situaciones de riesgo! Siga a rajatabla los calendarios de vacunación; tome medidas de higiene sexual tanto si no tiene pareja estable como si la tiene, y evite viajar a otras zonas del mundo sin tomar las precauciones básicas: vacunación si es necesaria, beber agua mineral, no comer nada en la calle.

C.G

Próxima semana
Todo lo que debe saber del cerebro

Appendix A4
Text 4: Short Enhanced Text

Medicina preventiva

Consejos útiles para llevar una vida saludable

Desde luego, tiene que comer bien. Acostúmbrese a pensar antes de abrir la nevera: cada día **tome** fruta y verduras, carne o pescado, dos o tres huevos por semana, leche o queso a diario ...

1182

Entérese muy bien de sus necesidades alimenticias y también de las de su familia (sobre todo si tiene hijos pequeños). **Haga** una lista y **pegue** la lista en la puerta de la nevera. Y aunque de vez en cuando se puede permitir un capricho, **adquiera** zanahorias o incluso algún fruto seco mejor que bollos o pan.

El ejercicio físico es también imprescindible. No tiene que convertirse en una gran atleta; simplemente **suba** y **baje** las escaleras de su casa y vaya por todas partes a pie y su corazón tendrá la agilidad que necesita.

Ni beber, ni fumar. Sabe perfectamente cuáles son los hábitos que le perjudican, pero otros no son tan obvios. Como trasnochar a menudo. Aunque de vez en cuando sea necesario trabajar hasta altas horas de la noche o salir con los amigos para echar una canita al aire, **duerma** sin interrupciones entre 7 y 8 horas para que su organismo pueda recuperarse. Tampoco se exponga al sol durante demasiado tiempo sin protección, tanto en la piel como en los ojos. Otros hábitos importantes tienen que ver con su seguridad: **use** casco si va en moto o bici y **conduzca** siempre descansado y, mejor, sin música.

Puede evitar muchas enfermedades. En ocasiones, existen en las familias antecedentes de enfermedades a las que usted puede tener propensión. **Pregunte** a sus mayores la historia clínica y **comente** con su médico de cabecera si usted o alguno de los suyos podrían desarrollar alguna alteración en el futuro y qué medios podría poner para evitarlo. En cualquier caso, **obtenga** un chequeo bianual - anual a partir de los 35-40 años - le será de gran ayuda para detectar posibles anomalias antes de que representen un serio peligro. ¡Y desde luego **evite** las situaciones de riesgo! Siga a rajatabla los calendarios de vacunación; tome medidas de higiene sexual tanto si no tiene pareja estable como si la tiene, y evite viajar a otras zonas del mundo sin tomar las precauciones básicas: vacunación si es necesaria, beber agua mineral, no comer nada en la calle.

C.G

Próxima semana

Todo lo que debe saber del cerebro

1183

Appendix B1

Comprehension Task

Now see if you can answer the following questions based on the article *Consejos útiles para llevar una vida saludable*, (Useful tips for leading a healthy life.), **Answer in English and do not look back.**

1. The article recommends that you put up a list on the fridge. What *two* pieces of information does the list have on it?

a. _____

b. _____

2. The article recommends that you eat sometimes *two* things instead of bread or rolls. What are they?

a. _____

b. _____

3. The article suggests that you do *three* kinds of physical activity to keep healthy. Can you list them?

a. _____

b. _____

c. _____

4 - 5. The article mentions *four* activities that are not good for your health. Two are very obvious while the other two are not. What are they?

very obvious

a. _____

b. _____

1184

not very obvious

a. _____

b. _____

6. How many hours of sleep does the article suggest you get per night?

a. _____ hours

7. There are *three* recommendations provided in this article for your safety. Can you list them?

a. _____

b. _____

c. _____

8. According to the article, *in addition to your doctor*, whom should you also consult to avoid illnesses?

a. _____

9. What does the article recommend people over 35 years to do?

a. _____

10. According to the article, there are *three* risky situations you should avoid in order to be healthy. Can you list them?

a. _____

b. _____

c. _____

Now turn over.

185

Appendix B2

Multiple-Choice Recognition Assessment Task

Here are some tips or advice for a healthy life. For each sentence, circle the letter which grammatically provides the tip, piece of advice, or recommendation.

1. (Be accustomed) _____ a pensar antes de abrir la nevera.

A. Acostúmbrate B. Acostúmbrete
C. Acostúmbrese D. Acostúmbrase

2. (Take) _____ fruta y verduras, carne o pescado.

A. Tomas B. Tomes
C. Toma D. Tome

3. (Find out)_____ muy bien de sus necesidades alimenticias.

A. Entérate B. Entérase
C. Entérese D. Entérete

4. (Make) _____ una lista de sus necesidades alimenticias.

A. Hace B. Haz
C. Haces D. Haga

5. (Paste)_____ la lista a la puerta de la nevera.

A. Pegue B. Pega
C. Pegues D. Pegas

6. (Get)_____ zanahorias o incluso algún fruto seco.

A. Adquieres B. Adquiere
C. Adquieras D. Adquiera

7. (Climb) _____ las escaleras de su casa.

A. Subes B. Subas
C. Sube D. Suba

8. (Go down) _____ también las escaleras de su casa.

A. Baja
B. Baje
C. Bajas
D. Bajes

9. (Sleep) _____ sin interrupciones durante la noche.

A. Duerma
B. Duermes
C. Duermas
D. Duerme

10. (Use) _____ casco (helmet) si va en moto.

A. Uses
B. Usa
C. Use
D. Usas

11. (Drive) _____ siempre descansado, mejor, sin música.

A. Conduces
B. Conduzcas
C. Conduce
D. Conduzca

12. (Avoid) _____ muchas enfermedades.

A. Evita
B. Evite
C. Evitas
D. Evites

13. (Ask) _____ a sus mayores la historia clínica familiar.

A. Pregunta
B. Preguntes
C. Pregunte
D. Preguntas

14. (Discuss) _____ con su médico sobre algunas enfermedades posibles.

A. Comentas
B. Comente
C. Comentes
D. Comenta

15. (Obtain) _____ un chequeo bianual -anual a partir de 35-40 años.

A. Obtenga
B. Obtienes
C. Obtengas
D. Obtiene

1187

Author

RONALD P. LEOW, Associate Professor, Department of Spanish and Portuguese, Georgetown University, 403 ICC, 37th and O Sts., NW, Washington, DC 20057. (202) 687-5840, fax: (202) 687-7083, e-mail: RLEOW@guvax.georgetown.edu. Specializations: second/foreign language acquisition, language curriculum development, teacher education, reading.

1188

Applied Language Learning
1997, Vol. 8, No. 2, pp. 183-216

An Interactive Model of Vocabulary Acquisition

Swathi Vanniarajan
San José State University

In this article, an attempt has been made to propose what I prefer to call "An Interactive Model of Vocabulary Acquisition." The model is interdisciplinary. More importantly, it goes beyond the traditional studies which equate vocabulary acquisition with learning to match the form and the meaning. Though the model attempts to account for the acquisition of knowledge as well as the mental processes involved in vocabulary acquisition, the emphasis is much more on the characterization of mental processes than on any other aspects. An important feature of this model is that it is L1 (first language) and L2 (second language) neutral, and can account for both. One of the major claims made in this article is that acquisition of phonetic form, morphology, syntax, semantics, pragmatics, and world knowledge converge in the acquisition of vocabulary and that the depth of knowledge one has about a word depends on one's experience with the word which determines the richness of its connections with other words. The article also contains a section on the pedagogical implications of the model.

Research has consistently shown that comprehension difficulty has the highest correlation with vocabulary difficulty than with any other factor including syntax (Chall, 1958; Dale, 1965; Klare, 1963; Nation & Coady, 1988; Yorio, 1971). Yet there have been considerably fewer research studies in vocabulary acquisition in both L1 and L2 than in syntax or phonology acquisition.

The major reasons for the comparatively sparse number of research studies in vocabulary acquisition are two: (1) it is generally assumed

© *1997, Swathi Vanniarajan*

1189

that the study of vocabulary acquisition does not belong to the realm of linguistics compared to the study of acquisition of abstract linguistic structures and (2) (more importantly, in my opinion), the previous and current research studies in L2 somehow tend to be out of conformity with each other as well as with L1 research due to a lack of a coherent model (of vocabulary acquisition) and of a workable framework.

Unfortunately, whether in L1 or L2, as Beck and McKeown (1991) claim, "a coherent theory of vocabulary research cannot exist of itself but must draw from theories that concern the full range of language comprehension and production" (p. 790). The theory, furthermore, must also attempt to capture the mental processes involved in inferring the meanings of unknown words and "establishing relationships between concepts, organization of concepts, and expansion and refinement of knowledge about individual words" (Beck & McKeown, 1991, p. 790).

In this perspective, an attempt has been made in this article to propose what I prefer to call "an interactive model of vocabulary acquisition." The proposed model draws its energy heavily from the information-processing models in psychology (Anderson, 1980, 1983; Collins & Loftus, 1975; Craik & Tulving, 1975; Daneman, 1991; LaBerge & Samuels, 1974).

The article is divided into three sections. The first section attempts to describe what it means to know a word. The second section contains an overall view of the proposed model of vocabulary acquisition. The third section deals with the implications of the model for vocabulary instruction.

What It Means to Know a Word

To know a word is to know more than its meaning. For psychologists, linguists, and applied linguists, to know a word is to have a knowledge of its lexical features, phonetic form, syntactic category, spelling, hyponyms, heteronyms, synonyms, antonyms, metanyms, metaphorical use, analogy, word-play, overextensions, underextensions, and appropriate sociolinguistic connotations, including style, register, and dialect (Anderson & Freebody, 1981; Dromi, 1993; Ellis, 1994; Nation, 1990; Pinker, 1984). For cognitive psychologists working in semantic network and spreading activation models, knowledge of a word includes (1) a knowledge of the degree of its phonetic, phonological, morphological, and semantic relatedness with other words (Collins & Loftus, 1975; Collins & Quillian, 1969; Lindsay & Norman, 1972); (2) an ability to verbally define what it means both contextually and decontextually (Nagy & Herman, 1987;

Nagy, Herman, & Anderson, 1985); and (3) an automated skill to activate and appropriately use it in both comprehension and production in real time discourse (Daneman, 1987, 1991; Daneman & Carpenter, 1980, 1983).

In addition to these, Beck, Perfetti, and McKeown (1982) observe that one's knowledge of vocabulary, in terms of quality, can be distinguished at three levels: unknown, acquainted, and established. According to them, a student who has a firmly established knowledge of a word can be said to have a different vocabulary from that of another student who has only an acquainted knowledge of the same word. Kameenui, Dixon, and Carnine (1987) also note that "it is conceivable that two students might know the same number of words and, possibly, even roughly the same words, but that those students have 'different vocabularies,' due to differences in the quality, or extensiveness of their knowledge of particular words" (p. 133).

The general claim is that one does not acquire a word instantaneously and that there are stages of acquisition between the firmly established and recently acquainted vocabulary. What this implies is that the complete acquisition of vocabulary in a learner is a time-consuming process.

An advantage of the recognition of this time consuming process and the stages in vocabulary acquisition is that it can account for why there exist two kinds of vocabulary in all of us, the kind that we can recognize while reading or listening but that we don't use in contrast with the kind that we can produce and recognize as well. Nation (1990) calls these "receptive" and "productive" uses of vocabulary, respectively.

Carey (1978) suggests that the stages of acquisition of vocabulary start with what she calls "fast mapping." Mapping, in Carey's sense, is a kind of abstraction. Fast mapping is what one does in his/her first encounter with the word. Roughly speaking, fast mapping may or may not capture all of the properties of the word. For a learner with low proficiency in the language, such an encounter may provide only an abstraction of only the proto-meaning of the word. Carey furthermore observes that after fast mapping, one needs "protracted further experiences" (p. 274) to complete the learning of the full sense of the meaning of the word. This she calls "extended mapping."

Carey also observes that extended mapping is a long-time process. With some words, it can even be a life-long process. In summary, knowledge of a word is much more than a knowledge of its meaning and acquisition of vocabulary is a cognitive process consisting of many stages. The process does not generally happen all at once.

What follows is a description of the interactive model of vocabulary acquisition. The goal is to account for the cognitive processes and the cognitive stages involved in it.

An Interactive Model of Vocabulary Acquisition

Interactive processing, in human information processing research, generally refers to the process of interaction between bottom-up (text-driven) and top-down (conceptually-driven) processings. In the present model, the word "interactive" is used in a different sense. It indicates the interaction between learner external (contextual) and learner internal (cognitive) factors. However, as in any other interactive model of human information processing, this model also ascribes a leading role to what is called "working memory." The overall framework of the model is given in Figure 1.

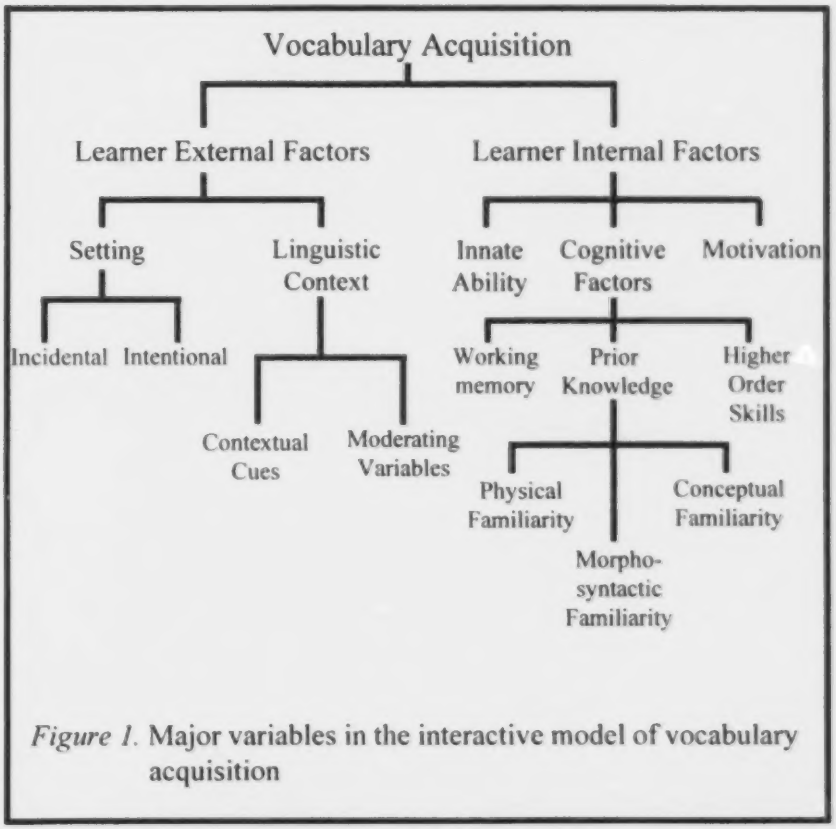


Figure 1. Major variables in the interactive model of vocabulary acquisition

Learner External Factors

Setting

The learner external factors which can affect the process of vocabulary acquisition are setting and context.

Setting can be divided into two types, incidental and intentional (Beck & McKeown, 1991; Ellis, 1994; Jenkins, Stein, & Wysocki, 1984; Nagy, Herman, & Anderson, 1985). In an incidental learning situation, which is considered to be the major mode of learning vocabulary, the intended goal is not the learning of words but some kind of active involvement with language. Such active involvement may include conversations, watching of movies or television programs, listening to classroom discourse, reading of magazines, books, brochures, etc. Intentional learning situations, on the other hand, refer to situations in which "the explicit purpose for an interaction is to learn the meaning of a word. These include a learner's decision to consult a source, such as a dictionary or more knowledgeable person, for direct instruction on the meanings of specific words" (Beck & McKeown, 1991, p. 798).

Experiments in L1 vocabulary acquisition on incidental and intentional settings (Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985) indicate that in incidental settings learners require at least six to ten exposures to the word before they can infer the complete meaning of the word. Nagy and Herman (1987), furthermore, suggest that every time the learners are exposed to the word, there is a small increase in their understanding of the meaning of the word. These findings support Carey (1978) that vocabulary acquisition starts with fast mapping followed by extended mappings.

Context

Context plays a major role in vocabulary acquisition. Beck, McKeown, and McCaslin (1983) propose that there may be a continuum of contexts with four identifiable points. They are: (1) misdirective contexts, (2) nondirective contexts, (3) general directive contexts, and (4) directive contexts.[1] Of the four, only directive contexts can lead learners, particularly L2 learners, to infer the correct meaning of the word through a supply of appropriate contextual clues.

Generally speaking, the contextual cues can be divided into two types: (1) linguistic cues and (2) nonlinguistic cues.

1193

Linguistic cues, in general, refer to the word's syntactic category, morphology, and case role. The effective use of linguistic cues depends on one's grammatical knowledge. Many research experiments have been conducted to find out how one's use of morphemic cues (knowledge of root word and any derivational or inflectional affixes attached to the word) can affect one's understanding of an unknown word.

For instance, in an experiment conducted by Sternberg, Powell, and Kaye (1983a), with L1 learners as subjects, the word thermoluminescence was used to explore how one might use the morphemic cues in meaning inference in unfamiliar vocabulary. One of their major findings was that learners who knew that thermo- is a prefix derived from the Greek root word therme, and that it refers to "heat," or that luminesce is derived from the Latin root word lumen meaning "light," were able to do far better than those who did not know these details. Their general finding was that learners use their structural, functional, morphological, and semantic knowledge of suffixes, prefixes, infixes, and root words to infer the meanings of unfamiliar vocabulary.

Nonlinguistic cues refer to the pragmatic hints contained in a discourse, which make possible the prediction of a word's meaning (Sternberg & Powell, 1983). There exist several classification schemes for various types of contextual cues (Ames, 1966; Dulin, 1969; Miller & Johnson-Laird, 1976; Sternberg, 1987; Sternberg & Powell, 1983; Sternberg, Powell, & Kaye, 1982, 1983a, 1983b). Of these, Sternberg (1987) is a good synthesis of the existing schemes. He proposes eight types of clues. These are: temporal cues, spatial cues, value cues, stative descriptive cues, functional descriptive cues, causal/enablement cues, class membership cues, and equivalence cues. Although no difficulty hierarchy has been set up among them, research in L1 vocabulary acquisition (Sternberg, 1987) indicates that of all the cues, equivalence cues, such as "a chrysanthemum is a flower," and descriptive cues, such as "the baklava tastes very sweet," are the simplest.

Sternberg (1987) furthermore suggests that the utilization of contextual cues is affected by what he calls "moderating variables." They are: (1) number of occurrences of the unknown word, (2) variability of contexts in which multiple occurrences of the unknown word appear, (3) importance of the unknown word to understanding the context in which it is embedded, (4) helpfulness of the surrounding context in understanding the meaning of the unknown word, (5) density of unknown words, and (6) usefulness of previously known information in cue utilization.

1194

In cognitive perspective, it can be seen that the first two variables are concerned with the provision of scope for extended mappings, since multiple occurrences (greater frequency) of a word result in multiple exposures. The use of the word in different contexts can, furthermore, increase the number of cues available, resulting in an increased probability that the meaning of the unknown word is inferred.

The next three variables are concerned with the relationship that exists among unknown words, the text in which they occur, and the learner. If the learner feels that the word unknown to him/her is necessary for understanding the surrounding material in which it is embedded, then s/he is likely to be strongly motivated to infer its meaning even in incidental settings. On the other hand, if the word is judged to be unimportant in the communicative context, then s/he is "unlikely to invest any great effort in figuring out what the word means" (Sternberg, 1987, p. 93). Also, if the cues are separated from the unknown word by a substantial portion of discourse, then the learner may fail to recognize the relevance of the cues. Indeed, the cues may be misinterpreted as relevant to some other unknown word to which they are proximal (Carnine, Kameenui, & Coyle, 1984).

Furthermore, whether the cues come before or after the unknown word may also play a crucial role in determining how far the cues can be helpful in inferring the meaning of the unknown word. Vanniarajan (1992) indicated that both native and nonnative speakers found pre-word cues more helpful than post-word cues.

A disproportionate occurrence of unknown words can also overwhelm the learner. Sternberg (1987) comments that "when the density of unknown words is high, it can be difficult to discern which of the cues that are available apply to which of the words that are unknown" (p. 92). In such contexts, the learner may be unwilling and/or may not be able to use the available cues to best advantage. This may have an effect on the processability of the text also.

The last of the moderating variables, usefulness of previously known information, is concerned with prior knowledge or what is called schema memory. Since prior knowledge is treated as a major variable under cognitive factors, it will not be discussed any further here.

Learner Internal Factors

While learner/external factors give us clues as to what is available for processing and in what degree of prominence, learner/internal factors provide us with insights into how the learner handles the available information. More specifically, they shed light on the kinds of cogni-

tive constraints the learner has to cope with while processing the available information. More importantly, the learner internal factors can also inform us why and how learners differ from one another in their vocabulary acquisition skill.

Learner internal factors can be divided into three parts: (1) the ability with which the learner comes to a learning situation and is considered to be innate (therefore universal), (2) those factors which are related to one's developing cognition (hereafter called cognitive factors: working memory, prior knowledge, and higher order skills), and (3) motivation.

Innate Ability

Knowledge of vocabulary is, indeed, encyclopedic. It comprises various kinds of knowledge about a word. Broadly speaking, they can be divided into three types: (1) physical knowledge or knowledge about the physical characteristics of the word—i.e., orthographic (spelling) and phonological (pronunciation) structure of the word; (2) morpho-syntactic knowledge, which subsumes the knowledge of what are called lexical features such as subcategorization; and (3) conceptual knowledge that includes semantic and pragmatic knowledge about the word.

Research in semantic network and spreading activation models, furthermore, suggests that knowledge of a word, comprises knowledge of how the word is related to other words and concepts (Collins & Loftus, 1975; Collins & Quillian, 1969).

In order to account for the three kinds of knowledge, as well as the relational knowledge, this paper assumes that human beings are born with an ability to abstract the physical, morpho-syntactic, and conceptual (including semantic-pragmatic) features of a word and store them in a kind of relational network consistent with the previously known words.[2] Figure 2 shows the kinds of abstractions of various types of knowledge made while encountering a word and the possible relationships that exist between words in our knowledge of vocabulary.

1196

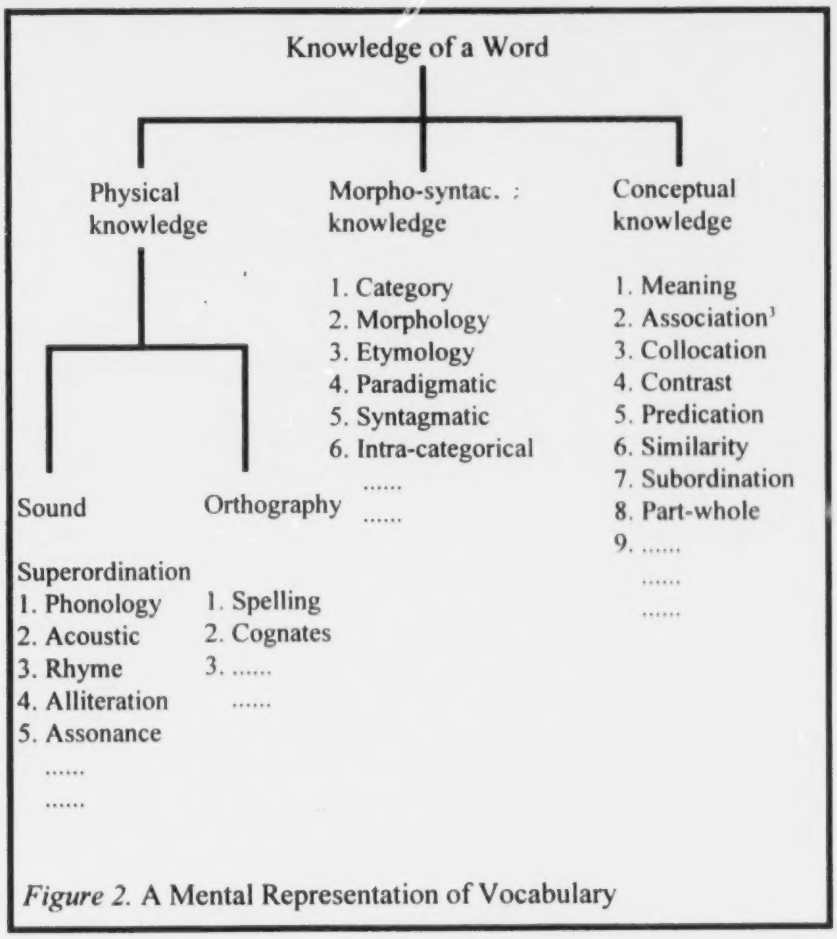


Figure 2. A Mental Representation of Vocabulary

The mental representation model presented in Figure 2 has much in common with the spreading activation theory and semantic network models in psychology (Anderson, 1976, 1980, 1983; Collins & Loftus, 1975; Collins & Quillian, 1969; Miller & Fellbaum, 1991). The common assumption underlying these models is that human beings are born with a capacity to extract (abstract) and store the conceptual structure of a word in a kind of relational network in their long-term memory. As such, the proposed model of vocabulary acquisition assumes that the various underlying features of a concept, as well as the word it signifies, are stored mentally as a set of entries bearing relationships to one another at all three levels — physically, morpho-syntactically, and conceptually.

The major difference between the model presented here and the semantic network models is that while the model presented here attempts to account for the acquisition of both the form and the meaning

1197

in a word, the semantic network models account only for the comprehension of the meaning in a word.

Assumption of an ability to abstract is not new to acquisitional linguistics. Grimshaw (1981) assumes that we are born with an ability to abstract the syntactic category of a word when we come across it. Pinker (1990) adds that "if children assumed that semantic and syntactic categories are correlated in restricted ways in the early input, then they could use semantic properties of words and phrases as evidence that they belong to certain syntactic categories" (p. 222). These assumptions are supported by the empirical findings that first and second language learners do not generally assign words to wrong classes. Research furthermore indicates that lexical substitutions do not occur between (syntactic) categories.

Along these lines, the proposed interactive model of vocabulary acquisition also assumes that language acquisition starts with lexicon and various levels of syntactic representations are the result of projections from the lexicon as otherwise it will not be possible to account for the relationship between d-structure and s-structure. Note that such an assumption is in agreement with the psycholinguistic models of first language acquisition as proposed in Anderson (1983), MacWhinney (1978), MacWhinney and Anderson (1986), and Pinker (1989, 1990).

This article will not enter into the controversy of whether the abstraction of lexical entries is made possible through purely distributional discovery or by associating the semantic property of a word with its distributional class.[4] The argument is, rather, that if the ability to abstract and store lexical features is innate, then there is not a priority reason why other features also cannot be abstracted simultaneously.

For example, Reich (1986, p. 40) claims that in L1 acquisition "children who are learning highly inflected languages sometimes apply the adult inflections to their idiomorphs." Fromkin (1993) argues that "the choice of inappropriate lexical items may occur because synonyms, antonyms, and similar sounding words are stored in close proximity to a given target word and thus may be retrieved in error" (p. 285). Fromkin (1993, p. 286) adds that "some word substitutions occur which appear to be influenced by previous words in the string; that is, they appear to be horizontally or syntagmatically conditioned," as shown in the following examples:

sesame seed crackers ⟶ sesame street
chamber music ⟶ chamber maid
gave birth at midnight ⟶ midwife

Fromkin calls these lexical selection errors. It seems to me these are much more than errors. They indicate that there exist close-knit associative links between the words intended and the words retrieved, as otherwise it would not be possible to explain the occurrence of these errors. Therefore, the interactive model assumes that we not only abstract the lexical features of a word but also its other features, and then we store them in a relational way compatible with the formal and conceptual features of the words already stored in the network.

The interactive model furthermore assumes, along with the spreading activation based semantic network models (Collins & Loftus, 1975), that there are shorter and longer as well as stronger and weaker links among words. When the links are closer and stronger, the activation of one word can result in what is called priming. In other words, the retrieval of one word can bring along with it the word which is closely linked to it.[5] In other words, short and strong links are considered to have a capacity to prime as well as get primed. Depending on the nature of the features involved, the priming can be at any level—phonetic, phonological, syntactic, semantic, or pragmatic.

For example, the word "Clinton" shall prime either "Hillary" (pragmatic relationship) to a Clinton biographer, or the word "clinic" (phonetic relationship) to a medical student, or even the concept of heaviness—"ton" (semantic relationship)—to a scrabble player. Like the neural network models, this model also assumes that knowledge about a word lies in the ever increasing connections.

In summary, information about a word is stored in a kind of encyclopedic network consisting of interconnected nodes in syntagmatic (linear), paradigmatic (associative-contrastive), and hierarchical relationships as proposed in Collins and Loftus (1975).

In this context, it is natural for anyone to wonder about the representation of vocabulary in a bilingual. Although there are research studies which suggest that learners start with independent L1 and L2 repertoires of vocabulary, then move on to have a common repertoire and then again have independent repertoires (a kind of U-curved pattern as it happens in the acquisition of morphology), it is generally believed that L1 and L2 lexicon are interdependent. The interdependence of L1 and L2 is believed to be indicated by negative L1 transfer which results in what is called "interlexicon" (Kruse, Pankhurst, & Sharwood Smith, 1987).

In my view, both representations are unitary, and together form a common repertoire right from the start. Note that it is frequently observed in L2 vocabulary research that "forms in one language clearly evoke the corresponding related forms in the other language" (Meara,

199

1980, p. 232). Also, a unitary representation of bilingual/trilingual vocabulary is not only cognitively economical but can also account for conceptual confusions in related vocabulary such as fahren in German and travel in English, or in false cognates such as will in German and will in English. This kind of representation can also account for the generation of nonce loan words, intra-word code-switching, and intra-phrasal code-switching.[6]

It can be seen from the above that innate ability is concerned with the universal capabilities of learners. It therefore deals with the invariant aspects of vocabulary acquisition among individuals. The next section, "Cognitive Factors," attempts to explain why there are quantitative and qualitative differences among learners in vocabulary acquisition.

Cognitive Factors

Working Memory

The working memory models as proposed in Baddeley (1981a, 1981b, 1986, 1989, 1992) suggest that working memory is a system that temporarily holds and manipulates information. It is "more like a workbench where material is constantly being handled, combined, and transformed" (Matlin, 1994, p. 126). Information processing research suggests that quantitative and qualitative differences among learners in their vocabulary knowledge may be due to working memory and two of its related variables: spaced learning and depth of processing. Figure 3 offers a simplified form of working memory within the human memory system.

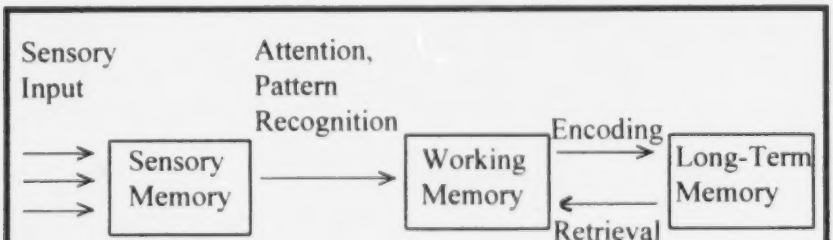


Figure 3. A Simplified Model of Human Memory (adapted from Atkinson and Shiffrin, 1968)

190

It can be seen from Figure 3 that working memory handles (online) information received from sensory registers and information retrieved from long-term memory. The current models of working memory are based on three important constructs: (1) it is a limited capacity attentional system, (2) it serves as a temporary store of the relevant information for later use in the performance of cognitive tasks, and (3) it has to be defined in functional rather than structural terms. Being a limited capacity attentional system, working memory allows an individual to attend to only a limited amount of information at one time. Being a temporary store of information, it enables an individual to store information temporarily for later computations. The functional rather than the structural definition of working memory emphasizes that the efficiency depends on the amount of combined energy required for processing and storage together. However, if there is a greater urgency for either of the two, then there will be a trade-off between them and all of the available energy will be used for one function, either for processing or for storage.

The relevance of working memory is seen especially in the comprehension of cohesive devices such as pronouns and contextual unknown vocabulary. Consider the following example from a psychology lecture in a university setting[7] to see the relevance of working memory in the computation of pronouns:

> Well, in that situation, what you will find is that the rat will press the—uh—lever. When the light is on, it will get food. It'll keep pressing the lever. When the light is off, it won't get food. Gradually it will discriminate between the light-on condition, light-off condition, and the response will take place only when the light is on, not when the light is off.

In the above example, the antecedent "rat" is mentioned only once at the beginning. However, there are four references to it in the form of the pronoun "it" in the utterances that follow. Unless the listeners have temporarily stored in their working memory what the antecedent is, they will not be able to figure out the references to "it."

The use of working memory in the comprehension of unknown vocabulary in contexts is also self-evident. When one comes across an unknown word, s/he has to process the relevant contextual cues, store them, integrate them and then compute the intended meaning of the word, while at the same time expending his/her mental energy on decoding (bottom-up processing) and comprehending the online discourse. This can be seen in the following example used in an

Y251 1201

experiment in Vanniarajan (1992):

> When Peter wanted to sue his doctor for malpractice, he was advised to consult and hire a nomiatrist, though he himself is a lawyer.

To infer the meaning of the word "nomiatrist"[8] in the above example, which is "a lawyer specializing in medical malpractice," one has to combine the contextual cues such as suing a doctor for malpractice and the hiring of a lawyer, as given in the linguistic cue (emphatic pronoun) "himself." In a real-time discourse, when processing (decoding) makes a considerable demand on working memory (as in the case of limited proficiency speakers), the individual may not be able to compute the entire meaning of the word[9] since he or she may not have any energy left for storage and integration processes, since the entire energy of the working memory may have been used up for processing alone. In such contexts, only fast mapping can take place.

Furthermore, in a real-life situation, one does not know when and where one may encounter a new word. Consequently, one does not generally keep the required energy on reserve to work on the contextual cues for meaning inference. Encountering an unknown word in this sense is an unwanted and unexpected expense which the budget does not take into account well in advance.

Craik and Lockhart (1972) argue that if an individual does not know beforehand what the intended task is, then his/her performance on that task must be termed incidental learning. By their argument, all unknown vocabulary encounters in a real-time online discourse are generally incidental learning situations. This does not mean that incidental learning situations do not involve attention. In fact, as Schmidt (1990, 1993) and Ellis (1994) suggest, incidental learning may require peripheral attention, if not focal attention.

Daneman and Carpenter (1980, 1983), LaBerge (1990), and Shiffrin and Schneider (1977, 1984) suggest that when capacity limitations exceed the combined processing and storage requirements of the task, then individuals can overcome the limitation by what is called "divide and conquer strategy." They can initially devote their entire attention to processing and then backtrack for working on the storage requirements, and then again to processing (i.e., integration of cues). Thus switching back and forth between the two, they can accomplish the task. For instance, an L2 learner who is at the initial stages of language learning may not be able to comprehend the meaning of the word in real-time reading. However, he or she may be able to

successfully infer the meaning of the word if given time to reread the materials. Support for this can be found in Vanniarajan (1992). In the study intended to compare the native and nonnative differences in inferring the meanings of unknown words in context, the researcher found that nonnative speakers generally were not able to infer the meanings of the words under timed conditions, but were able to significantly improve the quality of their definitions when given additional time. The findings furthermore indicated that the required time to complete the tasks also varied depending on the proficiency level and the background knowledge of the learner.

Overcoming the working memory constraints is, however, not an end itself, since Craik and Lockhart (1972) and Craik and Tulving (1975) suggest that it is one's depth of processing which actually determines how well the individual remembers the information. An important aspect of levels of processing theory is that it assumes various levels of information and that memory codes of each level are qualitatively different from one another. As for vocabulary in the proposed framework, it is assumed here that there are three distinct levels (see Figure 2): physical (phonetic, phonological, and spelling) at the top, morpho-syntactic in the middle, and conceptual (including semantic-pragmatic) at the bottom. It is also assumed here that each level, though it is distinct from others, subsumes its higher level.[10] For example, the conceptual level subsumes the morpho-syntactic and the physical levels, and the morpho-syntactic level subsumes the physical level. As such, when only the physical form of the word is analyzed, like its spelling or pronunciation, the encoded memory trace may turn out to be fragile and therefore may quickly fade away. On the other hand, if the analysis is at its deepest, which means that if attention has been given to the semantic, pragmatic, and the conceptual characteristics of the word, then the memory trace may last longer. Such memory traces are also likely to be represented in the vocabulary network in a way that is qualitatively richer.

Deeper analyses are likely to happen frequently in intentional settings, since they require a lot of attention. For Beck and McKeown (1991), "the theoretical underpinnings of the importance of deep processing during vocabulary instruction derive from the notion that instruction that requires the learner to actively generate information improves retention because it helps to build semantic network connections between new and prior information" (p. 806). In L2 vocabulary learning also, Cohen (1987) emphasizes the importance of deep processing and strongly advocates the incorporation of deep processing related vocabulary learning strategies such as mnemonic

1203

activities into L2 vocabulary instructional methods.

While deep processing has received some attention, the other psychological notion, "spaced learning," has not been adequately studied in both L1 and L2 vocabulary acquisition research. In fact, Pimsleur talks about this as early as 1967. In spite of his suggestion that repetitions must be increasingly spaced for better retention, somehow, the notion has not been given serious attention. According to the spaced learning theory, though deeper analysis can ensure that memory traces of words are encoded very deeply, yet if the vocabulary is left unused for long, it may decay, and eventually the individual may find it hard to retrieve it and may suffer from what is called "the transitional probability problem." This occurs because human memory is not like a tape recorder or the memory in a computer, wherein the stored memory can be retrieved anytime without affecting it in any way. Also, it is not mere reinforcement that decides whether an item is strengthened for further retrievals in one's memory. Recall the frequent claims made in L2 vocabulary research that reinforcement does not help students retain vocabulary (Mondria & Wit-De Boer, 1991).

According to the spaced reinforcement theory, the reinforcement must be increasingly spaced in the sense that reinforcement will have to be delayed to the extent that the recognition task is made almost like a recall task. In other words, once the learners learn certain vocabulary, they must be provided with delayed reinforcement rather than immediate enforcement. What this means is that learners must be required to recall and use the target vocabulary in a meaningful way on their own at a time they are about to find it difficult to do it. This may prove to be difficult for the learners. Yet the success of learning, in this case the successful retrieval of the learned vocabulary, will be greater because retrieval practice research suggests that the greater the difficulty with which learners retrieve the learned item, the more accessible it becomes for future retrievals (Jacoby, 1979).

Prior Knowledge

Following the mental representation of vocabulary framework (Figure 2), prior knowledge can also be divided into three types: physical familiarity, morpho-syntactic familiarity, and conceptual (semantic and pragmatic) familiarity. When the newly encountered word sounds like a familiar word or has the spelling of a familiar word, then it can be said that the word physically (orthographically and phonetically) resembles a known word. For example, cognates evoke physical familiarity. Research suggests that cognates with similar meanings are

1204

easy to learn. They are also called "true cognates." On the other hand, cognates which do not have similar meanings produce a negative transfer effect; they are, for this reason, called "false cognates." When a learner comes across a cognate while learning a second or foreign language, if the word has the same meaning, then it is possible that it may require minimal effort on the part of the learner to learn its meaning. On the other hand, if the word turns out to be a false cognate, then the learner may have problems inferring its meaning.

Morphological familiarity combined with the syntactic properties of their compositions can also help or affect one's learning of new vocabulary. Morphological familiarity refers to the morphological transparency in the word. There are two types. In one type, meanings can be inferred from the meanings of their components, and these are called morphologically transparent words: (e.g., "eyebrow.") In the other type, it seems, the meaning of the word can be inferred from its components though, in reality, it cannot be, and these are called deceptively transparent words: (e.g., "metaphysics.") Needless to say, true cognates and morphologically transparent words are easy to learn.

Conceptual familiarity refers to the semantic, pragmatic, and the world knowledge associated with the word. With regard to conceptual familiarity, closely following Drum and Konopak (1987), I propose that it is possible to classify individuals who come to the vocabulary learning situation in L2 into four types: (1) those who know neither the word nor the concept (new label, new concept), (2) those who have the concept but not the word (new label, old concept), (3) those who know the partial meaning (concept) of the word as he or she has a similar concept but not the same concept (new label, partial meaning), and (4) those who know a different meaning of the word (old label, new concept).

Of these four, the most frequently encountered problem in learning situations (more frequently in L2, though it is true of both L1 and L2) is acquiring the appropriate vocabulary for the already known concepts. For example, when a L2 learner comes across unknown vocabulary, if he or she has the concept, then he or she will be able to make an approximate guess of its meaning and override any processing difficulties. In other words, when one has the concept but not the word, then it is easy for him/her to predict the meaning of the unknown word. Research in L1 (Nagy, Anderson, & Herman, 1987) also suggests that incidental vocabulary acquisition takes place only for words for which the learners have concepts. A handicap of those who do not have the concepts is that they will have to learn both the meaning and the word. If the meaning happens to be cognitively too advanced or distant for

them, then they may learn only the physical structure of the word and not the meaning. In the case of those who have similar concepts, it is possible that the learner infers the word in a different sense (lack of precision) since the concepts are not always the same across languages. For those who already know the partial meaning of the word, it is likely that additional encounters result in extended mappings. When learners encounter polysemous words in a language, and if they do know a different meaning but not the intended meaning in the current context, then whether they will be able to learn the additional meaning of the word depends on whether they have the intended concept.

Higher Order Skills

Higher cognitive skills come by training and experience. The importance of these in language processing can be seen in claims that they correlate highly with one's language proficiency and can improve verbal ability. Dunmore (1989) claims that "the skill of using context to discover the meaning of unknown words is an effective aid to successful reading" (p. 338). Thorndike (1971) suggests that inferring the meaning of an unknown word from context is like solving a problem. It consists of selecting the right cues and putting them together in right relations. According to Sternberg and Powell (1983) and Sternberg (1987), there are at least three crucial processes which are applied to learning new words. They are: (1) selective encoding, (2) selective combination, and (3) selective comparison.

"Selective encoding involves separating relevant from irrelevant information for the purposes of formulating a definition" (Sternberg, 1987, p. 91). Hence, when we come across an unknown word, we analyze the available contextual cues in order to filter out the irrelevant details from further consideration. In other words, only the features of information relevant for the interpretation of the unknown word shall get activated and be held in working memory. Selective combination involves "combining relevant cues" into a single integrated memory representation "for a workable definition" (Sternberg, 1987, p. 91). The integration may be a combination of information held in working memory, and the related information retrieved from long-term memory (i.e., prior knowledge). "Selective comparison is a process by which new information about a word is related to old information already stored in memory" (Sternberg, 1987, p. 91). In other words, an interpretation is made out of integrated information.

For an example of how these processes are used, consider the following illustration given by Sternberg and Powell (1983, p. 885):

> Although for others, the party was a splendid success, the couple there on the blind date was not enjoying the festivities in the least. An acapnotic, he disliked her smoking; and when he removed his hat, she, who preferred 'ageless' men, eyed his increasing phalacrosis and grimaced.

The researchers point out that in the above passage:

> It is relevant to figuring out the meaning of "acapnotic" (a non-smoker) that he disliked her smoking (selective encoding - analysis) but it is not relevant that the couple was at a party. In figuring out the meaning of "phalacrosis" (baldness), one needs to combine at least four relevant cues (selective combination-integration) — that the woman likes "ageless" men, that the removal of the man's hat led to the conclusion that the man is not ageless, that the man with whom she has the date is showing increasing signs of phalacrosis, and that she grimaced as a result. Also, one must know that, among men, baldness often increases with age (prior knowledge), and even that baldness becomes visible when one removes one's hat (selective comparison-interpretation)(p. 885).

In their definitional studies conducted with bilingual children studying in the United Nations school in Manhattan, Snow, Cancino, Temple, and Schley (1991) suggest that "although skills already acquired in a first language can ultimately be transferred to a second language, this transfer evidently does not occur until a relatively high level" (p.109). One reason for this is that vocabulary acquisition skill requires the simultaneous operation of higher order skills while processing language. It is evident from research in working memory that only the advanced learners will have sufficient energy left over for higher order skills after meeting the processing demands (Daneman, 1987).

However, not all words may require higher order cognitive skills for the inference of their meanings. Also, there may be a kind of hierarchy among the vocabulary types, certain types of vocabulary requiring minimal use of higher order cognitive skills and certain others requiring greater involvement. It is likely that those which demand a greater involvement of higher order skills draw a great deal of processing energy from working memory.

.105207

Motivation

Second language acquisition literature has identified three kinds of motivation: instrumental motivation, integrative motivation, and achievement motivation (Deci & Ryan, 1985; Gardner & Lambert, 1972; Gardner & Maclyntre, 1991; Lukamani, 1972). Irrespective of its kind, motivation as a variable has been shown to play a crucial role in both incidental and instructional settings (Gardner & MacIntyre, 1991).

In instructional vocabulary learning settings, when the learner comes across an unknown word, he or she is likely to be more inclined and motivated to learn of its meaning if he or she has to meet certain standards of performance, like passing examinations. Once the learner has achieved success at this, it may provide further motivation to learn more and more, and in the process, the learner may acquire the necessary vocabulary acquiring skill. In incidental settings, expected standards of performance are set by the learner him/herself. In this setting, it can be said that the rate of vocabulary acquisition and the amount of vocabulary acquired are directly proportional to the communicational needs of the learner, the length of time he or she wants to be in contact with the language, and his/her purposes for acquiring the new language. For a learner, like a tourist with minimal but dire survival needs, the initial progress can be expected to be fast, but it may slow down after some time and the pace may virtually come to an end after the learning of the minimal vocabulary required to survive in the community.

On the other hand, for a new young immigrant whose goal is to adopt the new culture and attain near-native competence in the dominant language of the newly adopted country, the pace may turn out to be faster for quite a long time and the amount of vocabulary intended to be learned may also be large.

Interactive Model of Vocabulary Acquisition

It can be seen that in the interactive model of vocabulary acquisition, all factors, both learner internal and learner external, determine to a large degree whether, and to what degree, a learner will be able to acquire the word. The process is interactive because all these factors are indispensable and interact with each other at the same time. However, of all the variables, working memory can be considered to be the most important, since it is the place where the nature and the quality of processing are determined.

For example, learner external factors determine the saliency and

208

the processability of the target vocabulary, and in this way, they can be said to interact with working memory by determining whether the input is going to be noticed by the mechanism. However, it is to be noted here that motivation or some other cognitive factors can also ultimately decide the degree of processability as well as the acquirability of the word.

Also note that if the learner has some sort of prior knowledge of the unknown vocabulary, then he or she may have to expend very little energy on processing, and in the process, he or she may have sufficient energy available for storing the contextual cues and applying the higher order skills. As a result, he or she may acquire the word with little effort.

In summary, how long the learner needs to acquire a word depends on both the learner external and the learner internal factors. The interactive process is either complete or incomplete, efficient or inefficient, instantaneous or prolonged, depending how strong and/or efficient the variables are in the interactive process and how they interact with one another as well as how they interface with the functional capacity of the learner's working memory system.

Pedagogical Implications

The model presented here makes the following recommendations for efficient vocabulary instruction in formal L2 learning situations:

1. Teachers must learn to be patient with their learners when they teach new vocabulary. It takes time for new words to get firmly established in learners and teachers must not expect instant results.

2. Teachers must realize that there are different levels of vocabulary knowledge. A learner may be able to recall only partial knowledge of the word. This does not mean that the learner is incapable of learning the word. In fact, it is an evidence that he or she is in the process of learning the word. Like interlanguage, interlexicon can also be claimed to provide a window on the vocabulary learning mechanisms of the learner.

3. Teachers must attempt to simulate both incidental and intentional learning conditions in their classes in order to be naturalistic. Most of one's vocabulary is learned in incidental learning conditions. Hence, making learners get trained in learning new vocabulary in incidental conditions must be one of the top priorities in the classroom. However, it is better to start with intentional learning conditions since the learners will learn how to learn. Once they are confident in their learning skills,

91209

the teacher can move on to the incidental learning conditions.

4. Learners must be trained to identify the important and relevant vocabulary from the unimportant, peripheral, and irrelevant vocabulary for efficient understanding of the ongoing discourse. In my observations, some learners tend to think that all words in the text are important and usually spend undue time in looking up the meanings of all the unknown words. These learners must be made to understand that there is a great deal of redundancy in the language as a system and that it is not necessary nor is it an efficient way of comprehending the overall meaning of the text.

5. Most learners think that their success in L2 learning is impeded by their lack of vocabulary knowledge and that vocabulary learning is a difficult task. They think they have to memorize the form and the meanings of words. These learners must be made to understand that vocabulary learning is a skill and that rather than learning the form and the meaning, all they have to learn is how to learn new vocabulary.

6. Teachers must be trained to provide multiple encounters with the same word in different contexts. Multiple exposures will enable extended mappings and can also ensure that learners continue to learn these vocabulary items.

7. Equally important is that learners are exposed to vocabulary on appropriately spaced occasions for effective vocabulary learning. Mere multiple exposures will not be sufficient, especially if they are not adequately spaced. For this to happen, teachers must constantly monitor their learners and study informally when the learners begin to find it difficult to recall and use the new vocabulary. At such times, the teacher must reintroduce the vocabulary in meaningful and recall-like rather than recognition-like exercises.

8. Teachers must provide the necessary contextual support for the activation of background knowledge in learners. Vanniarajan (1992) indicated that even native speakers were not able to infer the meanings and therefore come up with appropriate definitions of certain words when they did not have the necessary background knowledge. It is therefore necessary for teachers to evaluate as well as raise the background knowledge of the learners whenever new vocabulary is introduced.

9. As vocabulary is assumed to be encoded in a kind of network, teachers can make learners play vocabulary association games. For example, the teacher can provide the learners with certain words and then ask them to come up with as many words as they can think of in relation to these words. When learners come up with uncommon associations, the teacher can ask the learners to explain why they made

9210

these associations. Such exercises can also allow the teacher to become familiar with the background knowledge of the learners.

10. It is important that both teachers and learners are made to understand that there is a continuum of contexts ranging from misdirective to directive contexts and that depending on their nature, contexts may reveal little or a great deal of information about a word's meaning. If learners understand that not all contexts lead to the successful inference of meanings of unknown vocabulary, then they may not end up frustrated if they are not successful.

11. Teachers must train learners to identify and use contextual cues in guessing the meaning of an unknown word wherever possible. Li (1988) claims that "since adequate cues in context can relieve learners of English as a second language from the anxiety of unfamiliar words, it might follow that reasonably sufficient contextual cues should be provided in texts for second language learners, so that enough information can be created for them to play the 'psycholinguistic guessing game' (Goodman, 1983) if the game is part of the instructional goal" (p. 411).

12. Research must come up with a continuum of difficulty among the moderating variables, contextual cues, and linguistic cues. Materials can then be prepared and presented in the classrooms accordingly.

13. Learners need to be trained to process the language with little attention to bottom-up processing. One of my observations in my Academic English classes is that students often complain that they are slow readers and that they don't have sufficient time to do the tasks. Recall the model's claim that learners can use higher order skills only when there is sufficient energy available for these to operate in working memory. If all the available energy has already been spent on bottom-up processing (i.e., for decoding), then the learners may not have any energy left over for exercising their higher order skills. In these cases, it is not that learners lack higher order skills, but that they don't have the necessary working memory energy left over to use these skills. In my opinion, this problem can be overcome by training learners in speed-reading and speed-listening tasks.

14. Vocabulary assessment tasks must aim to be both process- and product-oriented. Currently, most of the exercises are product-oriented. This must change. Process-oriented exercises must be given more importance than the product-oriented ones. Exercises which require learners to identify the relevant contextual and linguistic cues which are necessary for inferring the meanings of the unknown words and exercises which require them to identify the different types of contexts, from non-directive to directive, must be given abundantly in order to

1211

give learners ample practice in getting trained in acquiring vocabulary acquisition skills. Additionally, learners must be trained to integrate the cues which they identify and come up with appropriate meanings. They must also be trained to encapsule the meanings in appropriate definitions. These tasks may initially prove to be difficult for the learners. However, they must be made to understand that the rewards will be more than worth the pain.

15. Teachers must be trained to introduce new vocabulary in reading tasks rather than in listening tasks. The advantage is that compared to listening, in reading, the learner can always backtrack and adopt the divide and conquer strategy to overcome his/her L2 specific working memory capacity limitations. Also, while listening, if the discourse is at a speed with which the learner cannot cope, then he or she is most likely to miss the target meaning.

16. Finally, in situations where contextual information does not provide adequate information in figuring out the meaning of the unknown word, learners will have to be trained to use dictionaries. For Miller (1985), dictionary definitions are often in difficult formats. They generally contain only decontextualized definitions. Also, they do not provide the necessary background experience (assumed prior knowledge), as Chall (1987) rightly claims "every dictionary fails to be an encyclopedia" and does not cover "a wide range of human experiences." Out of the myriad decontextualized definitions given for a word in the dictionary, one has to choose a definition fitting his/her background knowledge and then translate these decontextualized definitions into appropriate contextual meanings. In my opinion, there is no guarantee that learners will be able to do this successfully without training.

Conclusion

The major strength of the proposed model is that it is different from the traditional models which focus only on the learning of form-meaning relationship. Unlike them, this model has made an attempt to capture how vocabulary is represented in the human mind, how vocabulary items are interconnected to one another, how we acquire new vocabulary, how cognitive and situational constraints limit one's ability to process and acquire unknown vocabulary, and what the pedagogical implications are.

The present model, furthermore, is not to be construed as antithetical to modular models of syntax. In fact, it is my contention that the representation may be unitary, but the procedural operations may be

1212

modular. An advantage that comes with my model is that, unlike other models, it can account for various types of lexicon: semantic lexicon, phonological lexicon, orthographic lexicon, pragmatic lexicon, etc. The major claim is that acquisition of phonetic form, morphology, syntax, semantics, pragmatics, and world knowledge converge in the acquisition of vocabulary[11] in both L1 and L2 and that the depth of knowledge one has about a word depends on his/her experience with the word which determines the richness of its connections with other words.

With regard to the overall evaluation of the model as presented here, all that can be claimed at this stage is that it is somewhat similar to Schumann's Acculturation model of L2 acquisition in its framework. Both models attempt to be all-encompassing. In this respect, some may feel that the model, like Schumann's, may not be able to fulfil the requirements of falsifiability hypothesis. In my opinion, with the current technology and the available sophisticated techniques in information processing research, it is possible to test every claim made in this model. A description of these techniques is beyond the scope of this article. I intend to conclude, therefore, by claiming that the success of this model, as it is, depends on how well it can be translated into experimental procedures and thus be verified.

Notes

[1] Although contextual inference is considered to be one of the powerful techniques of vocabulary learning, research in experimental psychology has repeatedly shown that it is neither a necessary nor the only effective technique of vocabulary learning in L1 (Beck & McKeown, 1991). It has also been shown that though children and L2 learners in naturalistic contexts learn vocabulary mostly through oral interactions and listening, not all contexts lead to the possible inference of the meaning of vocabulary (Beck et al., 1983). In other words, "often, natural texts do not reveal sufficient clues" (Beck & McKeown, 1991, p. 809). As such, one area of potential research is on the kinds of contexts which generally lead to the inferencing of incorrect meanings.

[2] For Meara (1980), L2 vocabulary in a learner is "less well organized and less easily accessible than those in the mental lexicon of a native speaker" (p. 238). The reason may be due to the qualitative differences between the two groups in their networks. The experimental research studies in L2 vocabulary acquisition also suggest that the definitions provided in the L2 (weaker language) are less stereotyped than in the L1 (stronger or dominant language) indicating that the

1213

networks of the native speakers are richer in connections.

[3] The difference between collocation and association is that the former refers to words that typically occur together while the latter refers to the words that can be associated with a given word.

[4] There is controversy in L1 acquisition literature as to how one comes to learn of the features of the semantic and syntactic features of a word. For Brown (1957), the L1 learner does it through a process of discovery by associating meaning and form. For Grimshaw (1981), it is the innate ability of the L1 learner which makes it possible. Both believe that meaning and form are inseparable and that the acquisition of syntactic features is not possible without learning the meaning of the word. For this reason, both were called functionalists. For distributionalists, on the other hand, meaning being redundant, a learner acquires the syntactic properties of a word by observing the distributional criteria of a word (i.e., its ability to occur where and with which). The problem with the distributional approach is that it makes the acquisition of syntactic features language specific since syntax differs from language to language. Also, since the distributional features are dependent on co-occurrence conditions, one has to posit two levels—one at the level of morphology and the other at the level of syntax. It has been shown in experimental psychology with experiments based on artificial languages that dependency grammar works well at the morpheme level but not at the syntactic level.

[5] The relationship, though perhaps bi-directional, need not be equal. In other words, the word "hospital" may activate the semantically associated word "doctor," but the word "doctor" may activate the word "nurse" and not "hospital." It is possible that these relationships, in kind as well as in strength, vary across individuals.

[6] The lack of significant time difference in the recognition of intraword code-switching and intra-phrasal codeswitching is a strong indicator that the representation is an all unified system. Again, the representation may be unitary, but the language specific processes may be modular.

[7] The data were collected from a psychology lecture given by a professor at the University of Ottawa in Canada.

1914

[8] This is a real word. This was taken from Dickson (1982) Words, p. 141.

[9] In this connection, it is necessary to recognize words which have only a single concept in their definitions and others which have multiple concepts. For example, "assist" can be defined by using one concept ("help") whereas "blood" will have to be defined by using multiple concepts ("red liquid circulating in arteries and veins of higher animals"). In my opinion, multiple concept words may take longer to be comprehended compared to single concept words.

[10] Whereas my assumption is in line with the claims made in experimental psychology that syntax is less deep than semantics, and semantics is less deep than pragmatics, Gass (1988, pp. 205-206) makes an opposite assumption: "One factor which determines whether or not comprehended input becomes integrated is the level of analysis of the input which the learner achieves. For example, it is possible that an analysis at the level of meaning is not as useful for intake as an analysis made at the level of syntax." Though she makes the claim with regard to the acquisition of syntax, I think, with regard to vocabulary a deeper analysis may benefit the learners.

[11] The major difference between lexicon and vocabulary is that whereas lexicon is pragmatics neutral, vocabulary carries the features of world knowledge as well. In other words, lexicon is the same (lexical features are basics and are linguistically neutral) across individuals which makes for uniformity in language acquisition. Vocabulary, on the other hand, is pragmatically loaded, individualistic, and is a major contributor to one's communicative competence. It is for this reason, linguists generally talk about lexicon, and applied linguists talk only about vocabulary. Also, it is evident by now that in the proposed model a distinction has also been made between vocabulary acquisition and vocabulary comprehension. Vocabulary comprehension would refer to the comprehension of its meaning alone whereas acquisition would refer to meaning and form, as well as whatever encyclopedic knowledge associated with it.

References

Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, *2*, 57-82.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale,

1215

NJ: Erlbaum.

Anderson, J. R. (1980). *Cognitive psychology and its implications*. San Francisco : W.H. Freeman.

Anderson, J.R. (1983). *The Architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, R. C., and Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Atkinson, R.C., and Shiffrin, R.M. (1968). Human memory, a proposed system and its control processes. In K. Spence and J. Spence (Eds.), *The Psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 85-195). New York: Academic Press.

Baddeley, A. D. (1981a). The concept of working memory: A view of its current state and probable future development. *Cognition*, 10, 17-23.

Baddeley, A. D. (1981b). Reading and working memory. *Bulletin of the British Psychological Society, 35*, 414-417.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Clarendon Press.

Baddeley, A. (1989). The uses of working memory. In P.R. Solomon, G.R. Goethals, C.M. Kelley, C.M. & B.R. Stephens (Eds.), *Memory: Interdisciplinary approaches* (pp. 107-123). New York: Springer-Verlag.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556-559.

Beck, I. L., & McKeown, M.G. (1991). Conditions of vocabulary acquisition. In R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 789-814). New York: Longman.

Beck, I. L., McKeown, M.G. & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *Elementary School Journal, 83*, 177-181.

Beck, I. L., Perfetti, C.A. & McKeown, M.G. (1982). The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*, 506-521.

Brown, R. (1957). Linguistic determinism and parts of speech. *Journal of Abnormal and Social Psychology, 55*, 1-5.

Carey, S. (1978). The child as word learner. In M. Halle, J.Bresnen, & G. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 265-293). Cambridge, MA: MIT Press.

Carnine, D., Kameenui, E. J. & Coyle, G. (1984). Utilization of contextual information in determining the meaning of unfamiliar words. *Reading Research Quarterly, 19*, 188-204.

T216

Chall, J. S. (1958). Readability: *An appraisal of research and application.* Columbus, OH: University Press.

Chall, J. S. (1987). Two vocabularies for reading: Recognition and meaning. In M.G. McKeown & M.E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 7-17). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, A. D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition, 9*, 43-62.

Collins, A. M., and Loftus, E. F. (1975). Spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 254* (11), 671-684.

Craik, F. I. M., and Lockhart, R. S. (1986). CHARM is not enough; Comments on Eich's model of cued recall. *Psychological Review, 93*, 360-364.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 248* (8), 240-248.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268-294.

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 895-901.

Daneman, M. (1987). Reading and working memory. In J.R. Beech & A.M. Colley (Eds.), *Cognitive approaches to reading* (pp. 57-86). New York: John Wiley and Sons.

Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 512-538). New York: Longman.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Daneman, M., & Carpenter, P.A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 561-583.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum.

Dickson, P. (1982). *Words*. New York: Delacorte Press.

Dromi, E. (1993). The mysteries of early lexical development:

1217

Underlying cognitive and linguistic processes in meaning acquisition. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 32-60). Norwood, NJ: Ablex.

Dulin, K. L. (1969). Using context clues in word recognition and comprehension. *The Reading Teacher, 23,* 440-445.

Drum, P. A. (1983). Vocabulary knowledge. In J. A. Niles & L. A. Harris (Eds.), *Searches for meaning in reading/language processes and instruction* (pp. 163-171). Rochester, NY: National Reading Conference.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73-87). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Dunmore, D. (1989). Using contextual clues to infer word meaning: an evaluation of current exercise types. *Reading in a Foreign Language, 6*(1), 337-347.

Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: A review essay. *Applied Language Learning, 5*, 1-32.

Elshout-Mohr, M., and van Daalen-Kapteijns, M. M. (1987). Cognitive processes in learning word meanings. In M. G. McKeown and M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 53-71). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Fromkin, V. (1993). Speech production. In J. B. Gleason & N. B. Ratner (Eds.), *Psycholinguistics* (pp. 272-301). Orlando, FL: Holt, Rinehart and Winston.

Gardner, R. C., & Lambert, W. E. (1972). *Attitude and motivation in second language learning*. Rowley, MA: Newbury House.

Gardner, R. C., & Maclyntre, P. D. (1991). An instrumental motivation in language study: Who says it isn't effective? *Studies in Second Language Acquisition, 13*, 57-72.

Gass, S. M. (1987). Introduction. *Studies in Second Language Acquisition* (Special Issue on the Second Language Lexicon), *9*, 129-132.

Goodman, K. (1983). Reading: A psycholinguistic guessing game. In J. T. Guthrie (Ed.). *Comprehension and teaching research reviews.* Newark, NJ: IRA.

Grimshaw, J. (1981). Subcategorization and grammatical relations. In A. Zaenen (Ed.). *Subjects and other subjects*. Indiana University Linguistics Club.

1218

Jenkins, J. R., Stein, M. L. & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21*(4), 767-787.

Kameenúi, E. J., Dixon, R. C., & Carnine, D. W. (1987). Issues in the design of vocabulary instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum Associates.

Klare, G. R. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.

Kruse, H., Pankhurst, J., & Sharwood Smith, M. (1987). A multiple word association probe in second language acquisition research. *Studies in Second Language Acquisition, 9*, 141-154.

LaBerge, D. L. (1990). Attention. *Psychological Science, 1,* 156-162.

LaBerge, D. L., & Samuels, S.J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Li, X. (1988). Effects of contextual cues on inferring and remembering meanings of new words. Applied Linguistics, 9(4), 402-413.

Lindsay, P.H., & Norman, D.A. (1972). *Human information processing: An introduction to psychology.* New York: Academic Press.

Lukamani, Y. (1972). Motivation to learn and language proficiency. *Language Learning, 22*, 262-273.

MacWhinney, B. (1978). The acquisition of morphophonology. Monographs of the Society for Research in Child Development, *43*, whole No.1.

Matlin, M. W. (1994). Cognition. New York: Harcourt Brace.

McKeown, M. G., & Curtis, M. E. (Eds.). (1987). *The nature of vocabulary acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language teaching and linguistics, 13*, 221-246.

Miller, G. A. (1985). Dictionaries of the mind. *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguists* (pp. 305-314). Chicago: Author.

Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. In B. Levin & S. Pinker (Eds.), *Lexical and conceptual semantics* (pp. 197-229). A Cognition Special Issue. Cambridge, MA: Blackwell.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

219

Mondria, J-A., and Wit-De Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics, 12*(3), 251-267.

Nagy, W. E., Anderson, R. & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*(2), 237-270.

Nagy, W. E., and Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. In M.G. McKeown and M.E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nagy, W. E., Herman, P. A. and Anderson, R. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* Newbury House.

Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching*. Longman.

Olshtain, E. (1987). The acquisition of new word formation processes in second language acquisition. *Studies in Second Language Acquisition, 9*, 221-232.

Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *Studies in Second Language Acquisition, 9*, 201-220.

Pimsleur, P. (1967). A memory schedule. *Modern Language Journal, 51*(2), 73-75.

Pinker, S. (1984). *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 399-441). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pinker, S. (1989). Language acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 359-399). Cambridge, MA: the MIT Press.

Pinker, S. (1990). Language acquisition. In D.N. Osherson & H. Lasnik (Eds.), *Language: An invitation to cognitive science* (pp. 199-241). Cambridge, MA: The MIT Press.

Reich, P. (1986). *Language development.* Englewood Cliffs, NJ: Prentice Hall.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

1220

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited. *Psychological Review*, 91-269-276.

Snow, C. E., Cancino, H, Temple, J. D., & Schley, S. (1991). Giving formal definitions: A linguistic or metalinguistic skill? In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 90-112). Cambridge University Press.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 878-893.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. (1982). The nature of verbal comprehension. *Poetics, 11*, 155-187.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. (1983a). The nature of verbal comprehension. In A. C. Wilkinson (Ed.), *Communicating with computers in classrooms: Prospects for applied cognitive science* (pp. 121-143). New York : Academic Press.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. (1983b). Teaching vocabulary building skills: a contextual approach. In A. C. Wilkinson (Ed.), *Classroom computers and cognitive science*. Academic Press.

Thorndike, E. (1971). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8*, 323-332.

Vanniarajan, S. (1992). Working memory and native-nonnative differences in inferencing the meanings of unknown words in context: An experimental study in interactive model perspective. Unpublished Ph.D. dissertation, Department of Applied Linguistics, University of California, Los Angeles.

Yorio, C. A. (1971). Some sources of reading problems for foreign language learners. *Language Learning, 21*, 107-115.

Acknowledgements

I sincerely thank the two anonymous reviewers for their helpful comments and suggestions on the original version of this paper. I also

1221

wish to thank Craig B. Chaudron, Thom Huebner, Robert Hvitfeldt, Kumaravadivelu, Denise Murray, Patricia Nichols, and my mentors John H. Schumann and Mari B. Wesche for giving me the benefit of their criticism on the early versions of this paper. The errors in the paper, however, are mine.

Author

SWATHI VANNIARAJAN. Assistant Professor, Department of Linguistics and Language Development, San José State University, San José, CA 95192-0093. E-mail: swathi@pacbell.net
Specializations: second language acquisition, psycho-cognitive linguistics, construct validity issues in language testing, and research methods.

1222

Applied Language Learning
1997, Vol. 8, No. 2, pp. 217-235

Pronunciation Instruction for "Fossilized" Learners Can it Help?

Tracey M. Derwing
University of Alberta

Murray J. Munro
Simon Fraser University

Grace Wiebe
University of Alberta

Thirteen adult students of English as a Second Language (ESL) enrolled in a speaking improvement program read aloud a list of true and false sentences at the beginning (Time 1) and end (Time 2) of their twelve-week course. Their recorded utterances were used in a listening task in which 37 native speakers transcribed the utterances as a measure of intelligibility. The listeners were also asked to judge utterances from Time 1 and Time 2 for degree of general comprehensibility and accent. An analysis of the transcriptions indicated that the utterances recorded at Time 2 were more intelligible than those produced at Time 1. However, only the true sentences were rated as significantly less accented and more comprehensible at Time 2 than at Time 1. The authors discuss the role of novelty in the processing of nonnative pronunciations. They also outline the nature of the speaking improvement program, which focused on general speaking habits as opposed to a concentration on individual segments.

Despite a renewed interest in pronunciation instruction (Morley, 1991), there is still very little empirical evidence available to guide teachers' choices of activities in the ESL classroom; nor is there much

© *1997, Tracey M. Derwing, Murray J. Munro, and Grace Wiebe*

1223

indication that classroom pronunciation instruction is effective. Although many descriptions of pronunciation practices and techniques have appeared in professional journals and there are several new classroom textbooks available, only a few studies actually address instructional efficacy. In view of the increase in the number of "accent reduction" and pronunciation improvement programs, and more focused attention on pronunciation within general ESL programs, we felt that an investigation of the effects of such instruction was warranted. We were particularly interested in determining whether the speech of long-term residents in an English-speaking environment would change noticeably, since many accent-specific programs are geared to people who are advanced speakers of English, long out of ESL classes.

Purcell and Suter (1980), in a reexamination of Suter (1976), failed to find a relationship between amount of pronunciation instruction (and for that matter, formal classroom ESL instruction) and pronunciation accuracy. They concluded that "attainment of accurate pronunciation in a second language is a matter substantially beyond the control of educators" (p. 286). A problem with this assessment, however, as Suter (1976) acknowledges, is that there was no indication of the quality or type of pronunciation instruction the students received, whether the students were motivated, whether the instruction was offered by native speakers (NSs), etc. Also, not enough background was provided about the nonnative speakers (NNSs) to know whether those who received pronunciation training were significantly less intelligible at the outset than those who did not. It is conceivable that specialized instruction brought those individuals into the normal range, but not to a point where their accuracy exceeded that of those who did not require instruction in the first place.

Perlmutter (1989), on the other hand, reported a significant improvement in second language (L2) speakers' intelligibility following a course which included extensive pronunciation instruction. The participants in Perlmutter's study were newcomers to the USA who were going to become teaching assistants at a university. Although it is encouraging that these individuals improved in intelligibility, it should be noted that learners typically display rapid progress during their first year in an English-speaking environment (Flege, 1988). Consequently, one cannot be certain that any of the improvement was due to the instruction itself.

Macdonald, Yule, and Powers (1994) undertook a comparative study of three common pedagogical approaches to pronunciation instruction: teacher-led drill, self-study in a language laboratory, and modified interaction, which they contrasted with a no-treatment

1224

condition. They taped 23 Mandarin speakers delivering a mini-lecture in each condition prior to intervention, immediately after intervention, and again, two days later. NSs then made forced-choice accentedness judgements where they were given two stimuli and were asked to select the one that was more native-like. There was a tremendous amount of individual variation in the NNSs' performance, so much so that only one significant difference of the students' pronunciation in the language lab condition was perceived to be significantly better than that of the students in the control condition immediately after intervention. As the authors point out, there are a number of limitations to this study, particularly the fact that the final effect of intervention was measured after only two days. Indeed, given that there was no sustained instruction, only two conditions mirrored what generally happens in some ESL programs: the control condition, in which students receive no pronunciation instruction, and the modified interaction condition, where teachers may incidentally ask for clarification when a student's pronunciation interferes with the intended meaning. Programs which offer accent-specific courses, either in the classroom or the language laboratory, usually provide sustained instruction over a period of several weeks. Notwithstanding the other problems cited by Macdonald et al. (1994) (e.g., the participants were all from a single linguistic group, there were few people in each condition, more time for intervention was allotted to the students in the language lab condition than the other conditions, etc.), the study is important in that it evaluated the efficacy of pronunciation instruction in terms of NS judgements of accentedness.

Accent judgements are a common measure of proficiency in pronunciation, but it has been argued that a better assessment of an individual's communicative needs may be obtained through intelligibility scores. Munro and Derwing (1995a, 1995b) and Derwing and Munro (1997) have shown that L2 learners who have intelligibility problems are almost always rated as having heavy accents. However, the reverse is not necessarily true. In fact, it is often the case that people who are judged to have a strong accent are perfectly intelligible, even though more processing time is sometimes required when NSs listen to accented speech (Munro & Derwing, 1995b). The additional processing time may explain why some utterances which are perfectly understood by listeners are nonetheless perceived to be somewhat difficult to understand. We define perception of the ease or difficulty of understanding as "comprehensibility."

In summary, it has not been established empirically whether a significant improvement in pronunciation can be achieved over a period

of several weeks in individuals who are not relative newcomers to an English-speaking environment. (Acton, 1984, has suggested an approach which he reports is successful in changing fossilized pronunciation, but he did not measure the change in any way.) In this experiment we assess NNSs' pronunciation at the beginning and end of a twelve-week speaking improvement course with measures of intelligibility, perceived comprehensibility, and perceived accentedness. Prior to the study, the participants had lived in English-speaking Canada for an average of ten years, and had all reported extensive daily interaction with native speakers, so they would not be expected to show noticeable progress over twelve weeks in the absence of instruction. In fact, one might not expect attendance at two evening classes a week to have more than a minimal impact on pronunciation skills. We therefore held very modest expectations for improvement in the group under investigation. If significant improvement appeared in any of the three measures, this would be directly attributable to instruction.

Procedure

Course Information

The two ESL instructors who offered the speech improvement course were asked to keep track of the course activities in notebooks provided by the researchers. At the end of the course the instructors gave their lesson plans and notes to us; they were also interviewed informally about the content of the course and their perceptions of the students' progress.

The instructors' course notes indicated that they focused on general speaking improvement, using what Firth (1992) has called the "zoom principle" (p. 173), that is, starting with overarching factors (such as prosodic features) which affect intelligibility, and gradually progressing to more localized factors. The teachers began the course with an explanation of the various contributors to accentedness; they then asked the learners to work in groups to determine what their own perceived needs were. They worked on body language, voice quality, volume (loudness), rate, and discourse markers; in addition to stress, intonation, and rhythm. Very little emphasis was placed on individual segments, in part because the students came from such diverse linguistic backgrounds, and in part because they appeared to have problems at a more global level, including some entrenched grammar errors which the teachers felt had an effect on intelligibility. The students did a great deal of group work in which they assessed each other's clarity; in

addition to teacher-developed activities, the classes used material from *Speaking Solutions* (Matthews, 1994) and *Clear Speech* (Gilbert, 1993).

Speaker Data

Thirteen NNSs, 10 females and 3 males, participated in this study; all had registered in a general speaking improvement course for high proficiency second-language learners. The first languages represented were as follows: 5 Mandarin, 2 Vietnamese, 2 Cantonese, 1 Farsi, 1 French, 1 Spanish, and 1 Ukrainian. The majority of the students self-selected the course, which was advertised for advanced speakers of English. (Four graduate students at the university were sent by their supervisor, who paid their tuition.) Those who self-selected reported a great deal of frustration with their accents; several individuals felt that their accents interfered with their performance at work. No proficiency tests were administered, but the three students who reported having taken a TOEFL in the past all scored 560 or higher. The students were drawn from two classes, and their teachers reported that they ranged from high intermediate to very advanced in proficiency. The average length of time they had lived in English-speaking Canada was ten years, with a range from two to twenty years. All participants had completed high school; some had postsecondary degrees.

Recordings/Digitizing

Individual recording sessions were conducted at the beginning of the course (Time 1) and twelve weeks later at the end of the course (Time 2). All sessions were held in a quiet room (for three evenings at Time 1 and two evenings at Time 2). The procedures at both times were identical, except that at Time 1 each student also completed a language background questionnaire. One of the researchers first asked the students to preview a list of true/false statements to identify any unfamiliar vocabulary, and provided explanations and modeling as required. False sentences were included to add a factor of unpredictability to the study. The sentences had been used in previous studies (Munro & Derwing, 1995b) and were found to be suitable for accent and comprehensibility judgements. (We hypothesized that the use of false sentences would limit the listeners' ability to construct meaning on the basis of their own knowledge of the world; that is, they would be fully reliant on the clarity of the individual's utterance.) The students were also shown a picture story and were told that they would be asked to describe the events depicted. They were encouraged to ask

for clarification if any pictures were confusing or if any new vocabulary was required. (The picture description narratives were collected for use in a separate study to be reported elsewhere.) When the students indicated that they were ready, they were shown into a second room where another researcher made the recordings. The students were seated in front of a Sony Cardioid F-VX300 microphone connected to a Marantz PMD201 cassette recorder. They read the sentences at a rate comfortable to them, and then, with the pictures in front of them, provided a narration of the story. Each session was 15 - 20 minutes long. The students were given an honorarium of $10 for each session. A total of 20 students participated at Time 1. Although all of these students continued to attend the course, only 13 returned at Time 2 (the last two classes) because of bitterly cold weather. For comparative purposes, parallel recordings were also made of four native speakers of English.

The recorded sentences were digitized and saved at 22kHz with 16-bit (CD quality) resolution, using a Macintosh computer. For the intelligibility task, one sentence was selected from each speaker at Time 1, and another was selected from Time 2. In each case a true and a false statement were selected from each speaker (7 true and 6 false at Time 1; 6 true and 7 false at Time 2). Although they were not included in the analyses, utterances from the four native speakers were also used, for a total of 30 stimuli: 13 sentences from Time 1, 13 from Time 2, and 4 native English utterances. The stimuli were re-recorded in random order on audiotape.

For the comprehensibility and accent judgement tasks, two identical sentences were selected from Time 1 and Time 2 for each of the 13 NNSs, for a total of 52 utterances, equally balanced between true and false. An additional 8 sentences were recorded by NSs as anchoring points for the listeners. These sentences were also digitized and rerecorded in a randomized order onto audiotape, giving a total of 60 stimuli.

Listener Data/Listening Sessions

Fifty-seven listeners were recruited from an undergraduate linguistics course as volunteers for the listening task. Of these, 37 reported on a language background questionnaire that they were monolingual native speakers of English; data from only this subset were used in the subsequent analyses. There were 34 females and 3 males, ranging in age from 19 - 44 years, with a mean of 24.6 years. All reported having normal hearing. A single listening session lasting less than an hour was held in a quiet room at the University of Alberta. (During the same

1228

session, data were also collected for another study.) The stimulus tape was played on a Sanyo MCD-Z31 stereo tape deck. Three examples were provided to give the participants practice with the transcription task and to establish a suitable volume level. The students were instructed to write, in standard orthography, exactly what they heard. One of the researchers controlled the rate of presentation so that a new stimulus item was not played until all listeners had completed the previous one. The comprehensibility and accent judgement tasks were done during a second round of listening. The participants were instructed to indicate on a 9-point scale the comprehensibility of each utterance (where 1 = extremely easy to understand and 9 = extremely difficult to understand). They then assessed degree of accentedness on a second 9-point scale (1= very little accent, 9 = very strong accent). Two example items were presented prior to the actual task.

Results

Intelligibility

The listeners' transcriptions were coded using an exact-word matching technique, which, although more conservative than other methods of scoring intelligibility, is simple and objective, and yields nearly identical results. With this technique, the transcriptions of the listeners are compared to the actual productions. The number of words exactly matching the original are counted for each utterance and converted to a percentage score for each listener.

Mean ratings for each listener were computed by calculating the average scores for the entire group of speakers. Data for only the non-native speakers were considered. A two-way repeated measures of analysis of variance (ANOVA) with time (before and after instruction) and truth value (true or false) revealed significant effects of time [$F(1, 36) = 37.52$, $p < .01$] and of truth value [$F(1, 36) = 18.60$, $p < .01$]. The two-way interaction failed to reach significance [$F(1, 36) = 2.74$, $p > .1$]. These findings indicate that there was a significant improvement in intelligibility scores from Time 1 (M=85%) to Time 2 (M=92%). It is interesting to note, however, that the overall intelligibility of the stimuli was affected by truth value, with true statements being more intelligible (M = 91%) than false utterances (M = 87%). Nonetheless, the lack of an interaction noted above indicates that the amount of improvement over time was about the same for the true and false statements. (See Figure 1)

1229

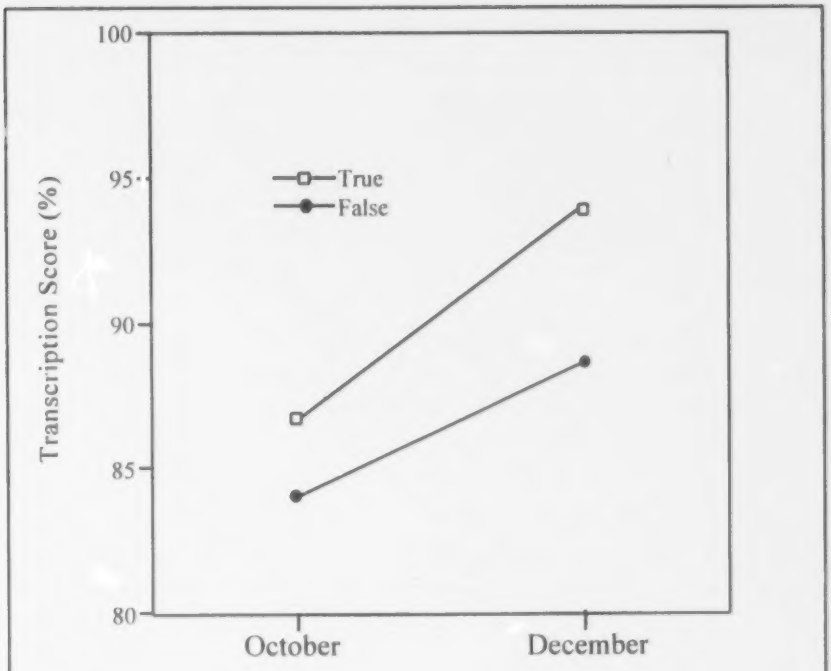


Figure 1. Improvement in intelligibility from Time 1 to Time 2.

Comprehensibility Judgements

A parallel analysis was carried out for the comprehensibility (ease of understanding) ratings. The two-way ANOVA revealed significant effects of time *[F* (1, 36) = 25.04, $p < .01$] and truth value [F(1, 36) = 15.96, $p < .01$], as well as a significant interaction [F(1, 36) = 12.43, $p < .01$]. An examination of simple main effects indicated that the interaction was due to a significant effect of time for the true statements but not for the false ones. In other words, as indicated in Figure 2, there was a significant improvement over time in the comprehensibility of the true statements only. (Note that a lower score indicates a higher level of comprehensibility.)

1230

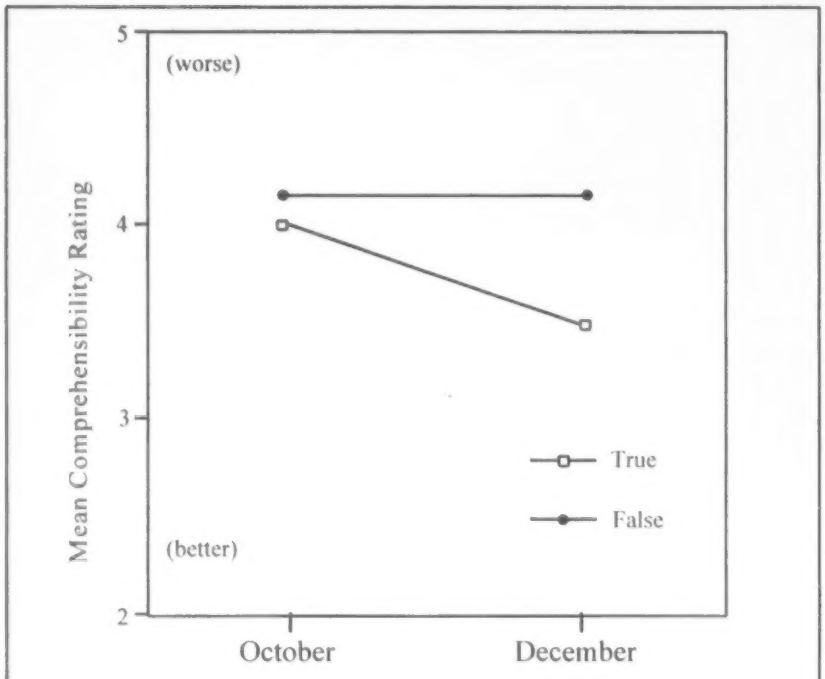


Figure 2. Improvement in comprehensibility from Time 1 to Time 2.

Accent Judgements

A third ANOVA, this time for the accent judgements, once again showed all three effects to be significant, that is, time [$F(1, 36) = 37.85$, $p < .01$], truth value [$F(1, 36) = 25.59$, $p < .01$] and their interaction [$F(1, 36) = 22.11$, $p < .01$]. Once again, true utterances were judged to be less accented at Time 2, but there was no significant change over time in the perception of the accentedness of the false statements (see Figure 3).

1231

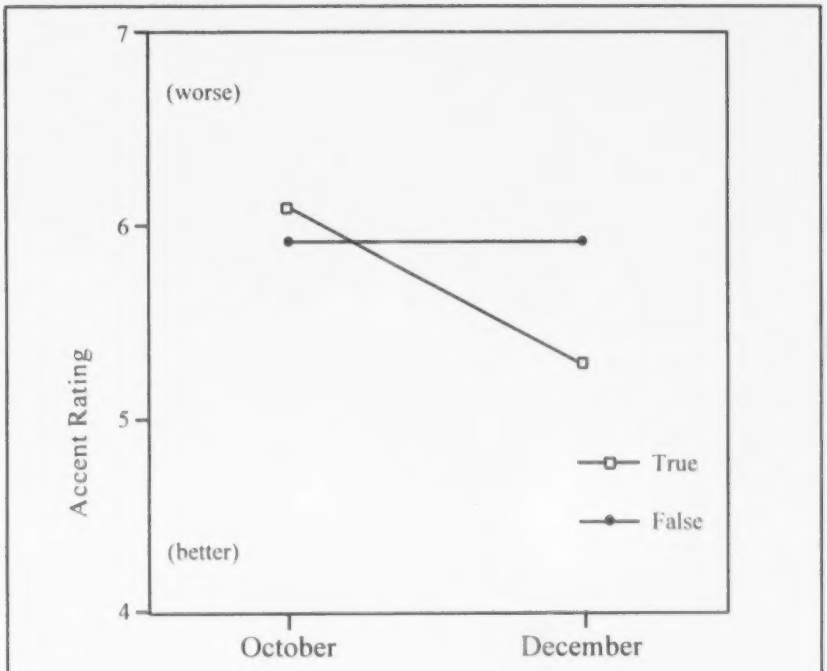


Figure 3. Improvement in accentedness from Time 1 to Time 2.

To compare these results with those of Munro and Derwing (1995a, 1995b), we identified the non-native utterances with mean intelligibility scores (pooled over listeners) greater than 98%. A histogram showing the distribution of all comprehensibility scores for these items (see Figure 4) exhibits a skewed distribution with a peak at 3. A small number of items received ratings greater than 5. A second histogram, this time for the accent scores (see Figure 5), reveals a noticeably different distribution. The peak in this case is at 7, and the ratings are more evenly distributed across the full range. However, none of the NNSs were identified as having very little accent (i.e., none were assigned a score of 1).

1232

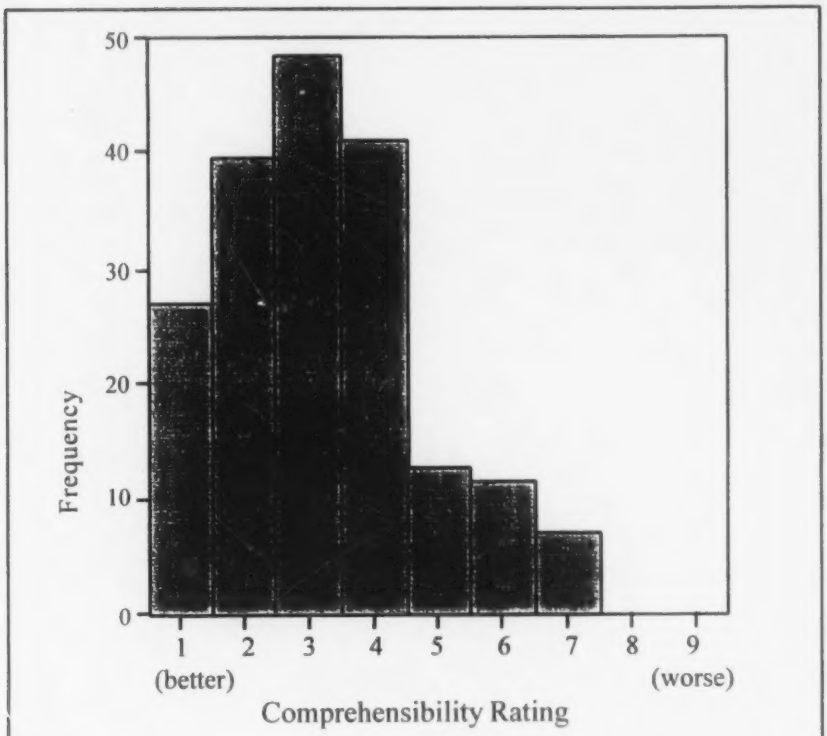


Figure 4. Comprehensibility ratings of sentences with highest intelligibility scores.

Individual Speaker Data

In addition to the analyses described above, we carried out t-tests on the intelligibility, comprehensibility, and accent scores for each speaker. To control for experiment-wise error, the p level was set at .004 (using the Bonferroni adjustment). The results of these tests are summarized in Table 1. Only three participants showed significant improvement on all three measures. However, eight individuals improved on at least one measure, while four individuals showed significantly weaker performance after instruction on one measure.

1233

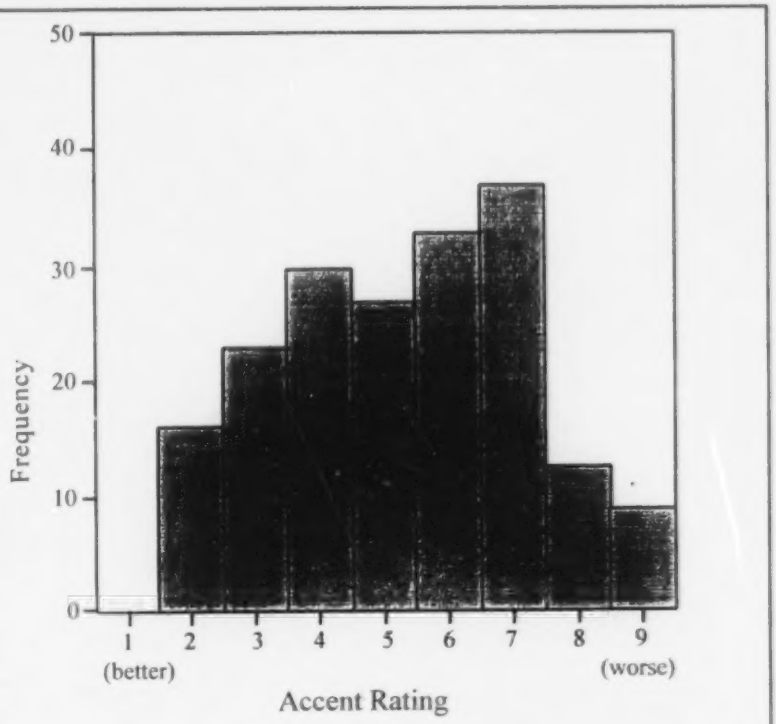


Figure 5. Accent ratings of sentences with highest intelligibility scores.

Teachers' Impressions

In post-course interviews, the instructors were convinced, first, that many of the students had improved, and second, that attitude and motivation were major factors in the success of the students. They reported seeing the strongest change in those individuals who consistently did the assigned homework and who came to class regularly. They also felt that those who self-selected to take the course were more motivated than the graduate students whose supervisor had sent them. The students and teachers alike reported that two three-hour classes a week for twelve weeks was too long, as most people's enthusiasm started to wane around the ninth week. Course participants agreed that a two-hour class twice a week for ten weeks would have been optimal.

1234

TABLE 1
Individual Speaker Summary

Student L1	Self-selected	Intelligibility	Comprehensibility	Accent
Cantonese	Y	-	n.c.	n.c.
Cantonese	Y	+	+	+
Farsi	N	n.c.	-	n.c.
French	Y	n.c.	n.c.	-
Mandarin	Y	n.c.	+	n.c.
Mandarin	N	+	+	+
Mandarin	Y	n.c.	+	+
Mandarin	N	n.c.	n.c.	+
Mandarin	Y	n.c.	n.c.	n.c.
Spanish	Y	+	+	+
Ukrainian	Y	+	-	n.c.
Vietnamese	Y	+	n.c.	n.c.

n.c. = no change
+ = significantly improved ($p < .004$)
- = significantly worse

Discussion

Intelligibility

The finding that listeners perceived an improvement in intelligibility across speakers over the twelve week period is particularly encouraging, given that the students registered in the course were not newcomers to Canada. Their mean length of stay was ten years, a point at which one would predict very little or no spontaneous improvement in a twelve week period. Although only three individuals improved significantly ($p<.004$), there was a trend in a positive direction for another five people; i.e., eight people's intelligibility scores at Time 2 were higher than those at Time 1. As Flege (1988) has shown, without instruction there is likely to be little spontaneous change in NNS productions after the first year in an L2 environment.

One possible explanation for the small number of significant improvements in intelligibility may be the existence of a ceiling effect. Four of the students who showed no significant change in intelligibil-

ity had a Time 1 score that was higher than the mean score for the whole group at Time 2 (Time 1 *M*= 85%; Time 2 *M*= 92%). Finally, some of the improvement noted by the teachers may have been related to aspects of their students' speech behavior that were not assessed in the listening tasks, such as body language and speech volume.

Comprehensibility and Accent Judgements

The relationship of the accent and comprehensibility judgements to intelligibility found here is very similar to that found in Munro and Derwing (1995a, 1995b) and Derwing and Munro (1997). In all four studies, accent judgements were harsher than comprehensibility judgements, and comprehensibility judgements indicated that listeners found some utterances hard to understand, despite the fact that they understood them perfectly. Although none of the participants in any of these studies received training or definitions from the researchers as to what to take into account when listening for accentedness, they seemed to be reliably capable of distinguishing accent from comprehensibility. In future research we plan to have listeners report which features they attend to for comparison with actual aspects of the NNSs' productions.

The before/after results in the comprehensibility and accent judgement tasks were very similar in that there was a significant improvement in the true sentences but no improvement in the false ones. (Both true and false sentences were included for every speaker at Time 1 and Time 2 in the listening tasks.) The distinction between the two sentence types is discussed in the next section.

In a few cases there was a deterioration of performance at Time 2 on one of the dimensions, and this finding demands some explanation. One possibility is the effect of such variables as physical or emotional state that might have had a detrimental influence on some individuals' performance. However, there is no reason to believe that such factors played anything more than a minor role in this experiment; if they had, one would have expected a balance of improvement and deterioration, but in fact the number of significant improvements far exceeded the incidences of deterioration, as can be seen by comparing the total number of pluses in Table 1 (16) with the total number of minuses (4).

Effect of Truth Value

Although it is unlikely that NNSs will produce many deliberately false utterances of the sort represented in this task in their day-to-day interactions, these sentences were included to simulate the unexpected,

novel quality of some NNSs' speech. Unusual lexical choice and ungrammaticality are both characteristic of some NNS discourse. Just as Gass and Varonis (1984) demonstrated that familiarity has a beneficial effect on intelligibility; its corollary, novelty, appears to have a deleterious effect (cf. Varonis and Gass's, 1982, finding that comprehensibility judgements can be influenced negatively by ungrammaticality). The presence of false novel utterances may reduce predictability in on-line processing, by reducing the effectiveness of the type of top-down processing that is required when working backwards through a semantically unclear utterance. Consider the following two sentences and their mean intelligibility scores produced by the same speaker: *Spaghetti grows on tall trees* (*M*= 83%) vs. *Ships travel on the water* (*M*=95%). In both these sentences, the first word was mispronounced; however, listeners had more difficulty with the former than with the latter on the intelligibility task. In the true utterance it was possible for listeners to misunderstand the word *ships*, yet still reconstruct it in the transcription task. The verb phrase in the false utterance, on the other hand, provides no clues as to the actual subject of the sentence.

Conclusion

This study has indicated that it is possible to destabilize pronunciation patterns in some individuals who have spent a long time in an English-speaking environment. In the absence of instruction one would not expect a significant spontaneous change in intelligibility or comprehensibility and accent scores over a twelve-week period. Despite our use of a conservative criterion for significance ($p < .004$), in just three months eight of the thirteen participants showed significant improvement on at least one of the three measures. In this case, an instructional approach which de-emphasized the importance of segmental units appeared to be effective (although the findings of this study do not provide contra-indications for a segmentally-based model).

There are a number of limitations here, including the fact that the listening tasks were carried out on isolated sentences. The use of sentences permitted the elicitation of judgement and intelligibility scores from a large number of listeners, while limiting the effects of familiarity; however, a discourse-based assessment of intelligibility, comprehensibility and accentedness would likely better reflect the students' performance in context (see Wennerstrom, 1994). Another limitation here was that only one method of instruction was employed; in a forthcoming study (Derwing, Munro, & Wiebe, 1997) we have extended

1237

the comparative approach of Macdonald et al. (1994) to a more realistic time frame than two days.

The role of motivation in relation to pronunciation instruction also requires further investigation. Three of the four graduate students who did not self-select to take the course were included in the study. Of those, one showed no improvement, one improved on the accent scale, and one improved significantly on all three measures. Although these people made it clear at the outset that they did not want to be in the course, we don't know to what extent their attitudes may have changed as the course went on. In future we would like to interview students regarding their motivation and monitor their effort as well as their progress over the duration of a speaking improvement course.

Finally, as Pennington and Richards (1986), Macdonald et al. (1994) and Morley (1991) have indicated, individual differences exist which cannot be traced directly to the first or second language influences. One size does not fit all; a closer examination of individual responses to pronunciation instruction might prove very revealing.

References

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18,* 71-85.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19,* 1-16.

Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1997). Global versus segmental approaches to the teaching of pronunciation: An empirical investigation. Manuscript submitted for publication.

Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American-English pronunciation* (pp. 173-183). Oxford: Oxford University Press.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84,* 70-79.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97,* 3125-3134.

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34,* 65-89.

Gilbert, J. B. (1993). *Clear speech* (2nd ed.). Cambridge:

1238

Cambridge University Press.

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve L2 pronunciation: The variable effects of different types of instruction. *Language Learning, 44,* 75-100.

Matthews, C. (1994). *Speaking solutions: Interaction, presentation, listening, and pronunciation skills*. Englewood Cliffs, NJ: Prentice Hall Regents.

Morley, J. (1991). The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly, 25,* 481-520.

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45*, 73-97.

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent and comprehensibility in the perception of native and foreign accented speech. *Language and Speech, 38*, 289-306.

Pennington, M. C., & Richards, J. (1986). Pronunciation revisited. *TESOL Quarterly, 20,* 207-26.

Perlmutter, M. (1989). Intelligibility rating of L2 speech pre- and post-intervention. *Perceptual and Motor Skills, 68*, 515-521.

Purcell, E., & Suter, R. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30,* 271-287.

Suter, R. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 30,* 271-287.

Varonis, E., & Gass, S. (1982). The comprehensibility of nonnative speech. *Studies in Second Language Acquisition, 4,* 114-136.

Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics, 15,* 399-420.

Appendix

True Sentences

1. Some roses have a beautiful smell.
2. Elephants are big animals.
3. You can buy a burger at McDonalds.
4. Hot and cold are opposites.
5. Gold is a valuable metal.
6. Exercise is good for your health.
7. Japan is a wealthy country.
8. Shakespeare wrote many fine plays.
9. Most teenagers like rock and roll.
10. Some people love to eat chocolate.

1239

11. Some people keep dogs as pets.
12. Young children can be very noisy.
13. Hungry cats like to chase mice.
14. You can start a fire with a match.
15. Red and green are colours.
16. Many houses are made of bricks.
17. Italy is a country in Europe.
18. Many people drink coffee for breakfast.
19. The American flag has stars and stripes.
20. Ships travel on the water.

False Sentences

21. You can buy beer at church.
22. Spaghetti grows on tall trees.
23. The Queen of England lives in Quebec.
24. Milk comes from yellow chickens.
25. Most swim suits have long sleeves.
26. A monkey is a kind of bird.
27. The sun always sets in the north.
28. The inside of an egg is blue.
29. August is a winter month.
30. People eat through their noses.
31. A nickel is worth twenty-five cents.
32. There are many cities on the moon.
33. Gasoline is an excellent drink.
34. It always snows in July.
35. March has thirty-eight days.
36. Most people wear hats on their feet.
37. The stars come out in the day.
38. Wednesday is the first day of the week.
39. All men can have babies.
40. All dogs have fifteen legs.

Acknowledgments

We express our appreciation to the participating students and their instructors, Liz Karra and Karen Carter of Grant MacEwan Community College, for their gracious cooperation. We also thank Kama Jamieson for her assistance throughout the project. The comments of Bruce Derwing and two anonymous reviewers were helpful. This study was

funded by a Support for the Advancement of Scholarship grant from the Faculty of Education, University of Alberta.

Authors

TRACEY M. DERWING, Associate Professor, Department of Educational Psychology, University of Alberta, Edmonton, AB, T6G 2G5, Canada, E-mail: tracey.derwing@ualberta.ca. Tel: (403) 492-3668. Specializations: NS-NNS interaction, accent and intelligibility, L2 literacy.

MURRAY J. MUNRO, Assistant Professor, Department of Linguistics, Simon Fraser University, Burnaby, B. C., V5A 1S6, Canada. E-mail: mjmunro@sfu.ca. Tel: (604) 492-8272. Specializations: experimental phonology, pedagogical development.

GRACE E. WIEBE, Visiting Assistant Professor, Department of Linguistics, University of Alberta, Edmonton, AB, T6G 2E7, Canada. E-mail: grace. wiebe@ualberta.ca. Tel: (403) 492-8272. Specializations: experimental phonology, pedagogical development.

1241

Applied Language Learning
1997, Vol. 8, No. 2, pp. 237-260

What Researchers Say and Practitioners Do
Perspectives on Conscious Grammar Instruction in the ESL Classroom

Miriam Eisenstein Ebsworth
New York University

C. William Schweers
University of Puerto Rico

In this article we report the approaches of researchers and methodologists to pedagogical grammar and the answers two groups of teachers have developed with respect to the role of explicit grammar instruction for adult English as a Second Language (ESL) learners. A review of suggestions from the literature is contrasted with the reported practice and beliefs of 60 college-level ESL teachers, 30 from Puerto Rico and 30 from the New York area.

Based on a questionnaire, interviews with a subset of the participants, and interpretive discussions with two experts, we found that a majority of the teachers from both groups had clearly defined approaches to grammar teaching. There were between and within group distinctions, however, which are explained by different contexts of teaching as well as individual differences in teacher approach.

ESL teachers have questions. It's natural for them to have questions. They question the techniques they use when interacting with students, they question the methodologies they employ, they question just about every aspect of the ESL teaching-learning process. One question in particular that teachers often raise is what the role of explicit grammar instruction in the ESL classroom might be, if indeed it has a role.

© 1997, Miriam Eisenstein Ebsworth and C. William Schweers

1242

Teachers also have answers. To be able to teach their daily classes, teachers have to come to some personal accommodation with respect to the many doubts they may have. Their responses come out of their years of teaching experience as well as their familiarity with the literature. Woods (1996) recognizes the importance of incorporating teachers' understanding of language learning in the evolution of theory and practice. In this paper, we will be looking at the approaches of researchers and methodologists and the answers teachers have come up with regarding the role of explicit grammar instruction for ESL learners. All these perspectives are necessary in our search to define grammar's role in language instruction and should be taken into account. Accordingly, we shall begin by reviewing the viewpoints of theoreticians and researchers with respect to this issue. Then we report and interpret the results of a study we conducted of teachers' opinions about grammar's role in language teaching and learning. The participants in the study were 60 university teachers of ESL from Puerto Rico and the New York area with a range of experience from three to forty-five years. Through a questionnaire and post-hoc interviews, we explored their opinions and practice regarding the role of pedagogical grammar in college ESL teaching.

A Brief Review of the Arguments

Traditionally grammar instruction was the heart and soul of the second language class. But with time, researchers and teachers began to question the utility of inundating students with grammar rules when they were seldom able to apply them to simple communication. This led to the development of communicative methodologies (Howatt, 1984). These were often premised on the assumption that comprehensible input would naturally lead to the acquisition of grammar rules without explicitly teaching them. It was felt that meaningful interaction under anxiety-free conditions was sufficient for language learning (Krashen,1994).The negotiation of meaning during communicative interaction has also been identified as a factor leading to naturalistic acquisition (Long & Crooks, 1992; Long & Porter, 1985; Pica, 1987; Wenden, 1991).

What was never specified, however, was the intensity of comprehensible input and meaningful interaction necessary to trigger natural acquisition mechanisms. In cases where students' only exposure to the language is a few hours of classroom instruction per week, communicative approaches proved to only be of limited value. Schulz (1991) observed that "time available in a conventional foreign language pro-

gram... is simply inadequate, if we hope to have them develop any meaningful, lasting communicative proficiency" (p. 22).

With respect to the recently perceived need to return to some sort of explicit grammar instruction, Fotos (1994) notes that

> A compelling body of evidence has accumulated recently supporting the position that formal instruction on language properties is related to the subsequent acquisition of those properties. These findings present a dilemma for many teachers who have become committed to the use of communicative approaches to language learning, wherein learners are given a rich variety of comprehensible input, and teacher-fronted grammar instruction is generally omitted (p. 323).

Based on observations such as this, researchers have begun once more to explore the role of explicit grammar instruction in language learning. These researchers have discovered a number of interesting things about the role of implicit instruction. It has been found that where natural acquisition principles divide the learning of a form into developmental stages, the passage through these stages can be accelerated through instruction. Where the grammar item to be learned can be hierarchized into implicational scales, it has been found that teaching the most difficult structure first speeds up the learning of the whole range of structures. Items which are "easy to learn" such as the plural *s* are learned more readily when taught explicitly than when acquired naturalistically. This is because features such as the *s* are often not salient in input. Grammar instruction also appears to be effective with those grammatical features which are so close to the learners' first language (L1) and/or so rare in the second language (L2) input that it is difficult for them to notice the difference between the two (Pica, 1994). Thus, according to some researchers, explicit instruction does seem to affect acquisition under certain conditions (N. Ellis, 1994; Robinson, 1997). To quote Teresa Pica (1994),

> The effectiveness of grammar instruction appears to depend largely on selection and sequencing of grammar rules and careful assessment of learner readiness. Some items are better off not taught, while the learning of others is enhanced, indeed accelerated through instruction (p. 67).

For these researchers, explicit grammar instruction does have a definite place in second language teaching and learning. The problem

1244

is [illegible]ing those grammar items which are teachable at a given point in learner development (Pienemann, 1984, 1987). In spite of these findings, other researchers still maintain that comprehensible input in a context of low anxiety is sufficient for language learning and that explicit teaching of grammar is superfluous. For example, Krashen (1982, 1985, 1994, 1996) and Krashen and Terrell (1983) strongly argue for this position. (Krashen's view will be elaborated later in this discussion.) Schulz (1991) also forcibly supports this point of view. She states that "Formal instruction (i.e., grammar analysis and discrete-point grammar practice) can temporarily improve performance on discrete-point tests, but apparently has relatively little influence on spontaneous language use" (p. 20). Furthermore, researchers such as Harley (1993) question the degree to which particular items can be successfully taught for particular learner populations at varying points of development.

Thus opinions range from a "zero position" with respect to the teaching of grammar to those who advocate a strong emphasis on grammar instruction. The mid-point, however, is perhaps the most commonly held stance. Most researchers recommend a communicative approach with consciousness raising activities focusing on grammar points introduced in the context of a communicative task (Long & Crooks, 1992; Montgomery & Eisenstein, 1986; Savignon, 1991; Fotos, 1994; Van Patten & Cadierno, 1993). Rutherford and Sharwood-Smith (1988) reflect a range of opinions on the appropriate use and status of grammar in the language classroom.

Several noted researchers have recognized the important role that teachers can play in exploring the potential of conscious grammar focus and instruction in L2 research and curriculum development. Nunan (1989, 1990) provides explicit guidelines to encourage and support practitioners in their research efforts. Such "action research" is also advocated by Williams (1995) in her recent article on the role of teachers in implementing and assessing a focus on form in communicative language teaching. As we try to unravel the varying opinions in the research literature with respect to explicitly focusing on grammar, we should remember Sharwood Smith's (1993) injunction that "It is still an open question as to how much conscious awareness of the formal properties of language, and hence instruction based on inducing this awareness, actually helps the development of spontaneous language use" (p. 172).

A Historical Perspective

In our attempt to find order in our communicative systems, we

245

have analyzed language into a number of overlapping subsystems. Primary among these, for teaching and learning purposes, are grammatical structure or syntax. While linguists explore the nature of syntactic structure, methodologists and teachers consider the role of explicit grammar instruction in enhancing language acquisition for learners.

Several texts review in depth the various swings of the pendulum regarding how our evolving knowledge of grammar can be used in language teaching (Howatt, 1984; Kelly, 1969; Stern, 1983). Questions range from the relationship of a learner's conscious grammatical knowledge with language production and comprehension to whether conscious grammar knowledge can be a potential barrier to fluent production. What is clear is that there is as yet no consensus regarding the appropriate role of conscious grammar learning in second language acquisition. These texts give us a sense of our vacillating journey regarding grammar and second language teaching.

According to Kelly (1969), the use of grammar explanations in education dates back at least to the ninth century. Until the invention of printing and books in the 5th century, grammar rules were taught orally, sometimes through verse. During the Renaissance, some teachers advocated the use of grammar books to avoid "barbarisms." Comenius, whose *Restauratio Linguae Atinae* was published in 1533, advocated completely abandoning grammar as a teaching method. Erasmus, in a 1520 letter, favored the use of formal grammar for beginners but believed that inductive methods were more useful at advanced levels. At Port Royale in the 17th century, "general grammar" learning was proposed; knowledge of L1 grammar was thought to be transferable to the second language, as languages had so much in common.

In the 19th century there was an interest in "natural" L2 learning. The natural methodologists of this time repudiated the usefulness of grammar teaching and focused on acquisition through conversation. It was this natural method that spawned the direct method used at the beginning of the 20th century and ushered in the audio-lingual era. Kelly, Howatt, and Stern's analyses of the range of views regarding language teaching mirror the importance given to social, philosophical, and literary aims of language teaching.

Current Perspectives

While communicative teaching has been the state of the art for some time (Littlewood, 1981; Richards and Rogers, 1986), the role of implicit and explicit grammar techniques remains unresolved. One

1246

popular and current view is that a pedagogical grammar can play an effective role in "consciousness raising" for the learner, and that this conscious identification of grammatical constructs in the target language will promote acquisition and learning; however, the term consciousness is not well understood.

Consciousness and Meta-Awareness

Rutherford and Sharwood Smith (1988, p. 3) note that "consciousness raising is intended to embrace a continuum ranging from intensive promotion of conscious awareness through pedagogical rule articulation on the one end to the mere exposure of the learner to specific grammatical phenomena on the other." Orse (1976) in Trampe (1994) sees consciousness as an expression of an individual's brain activity interacting with environment. Trampe (1994) agrees that consciousness is difficult to define. Schmidt (1990) makes the strong claim that conscious processing is needed as one step in the language learning process; he divides consciousness into several parts, one of which is awareness which includes perception, noticing, and understanding. Schmidt points out that for some, consciousness of grammar means that a learner can articulate a language rule or pattern while Chomsky claims that principles of grammar are inherent but (unconsciously) activated by language (Chomsky, 1986). Bialystok (1994) finds that consciousness in second language acquisition "tends to be associated with high-level limited capacity processing" (p.164). With such inconsistent characterizations, it is no wonder that educators find a consideration of the research on consciousness in grammar acquisition perplexing.

There is a rich literature regarding the possible roles of consciousness and meta-awareness (awareness of how language functions) with respect to information processing, storage and retrieval, both in general and with specific reference to language. In a 1979 paper, Bialystok considers the influence of formal explicit versus implicit knowledge in forming second language grammaticality judgments. Bialystok (1990) notes that "children progress by making explicit their knowledge of language" (p. 122). However, she goes on the ponder which comes first, the articulation of a language principle or consciousness of it. A potentially useful distinction made by theorists is the difference between declarative knowledge (about linguistic phenomena) and procedural knowledge (how to realize language operations) (Anderson, 1983). There is a range of opinions about which may come first and how the two are related. In question is the degree to which these

are two separate categories which develop independently or interdependently. McLaughlin (1990) explains how focused attention to aspects of L2 input can play a meaningful role in an information processing model.

Another useful addition to our evolving understanding of grammar instruction is suggested by Pennington (1995) whose collection of teacher-generated grammar teaching modules represents an underlying philosophy which she characterizes as, "a situated process view of grammar learning" (p. vii). This view transforms the student's role into that of a creative and active participant engaged in a process in which grammar options are matched to contextualized meaning with adequate opportunities for production and feedback. A related perspective is articulated by Batstone (1994) who emphasizes that a task-based approach can result in raising learners' consciousness of the processes of language use.

Views Denying a Role for Explicit Grammar Instruction

Krashen (1982, 1985, 1994,1996) states the strong theoretical position that comprehensible input at an appropriate developmental level and low anxiety (a low affective filter) are sufficient conditions for second language acquisition. He delegates grammar learning to a separate system requiring conscious attention and time for retrieval in order to monitor output. He is also among those who assume the influence of language universals on a natural order of learners' syntactic interlanguage development. While his theory of second language acquisition and learning undergoes constant evolution, the cornerstones remain. Central to his view is that while a learner's conscious attention is focused on meaning, her subconscious system accepts input which is available for analysis of the rule governed aspects of the second language. The result of this analysis becomes part of the learner's acquired system.

Second language methodologists have the challenging task of synthesizing our knowledge, philosophy, theory, and sometimes speculation about grammars and consciousness and translating it all into usable suggestions for language teachers. Krashen and Terrell, in the Natural Approach (1983) are among the strongest proponents of a methodology focused on meaningful comprehension and interaction with a very limited role for conscious grammatical knowledge (mostly acquired independently by interested learners using grammar texts).

An intermediate view is held by proponents of the Silent Way (Gattegno, 1976) who do not advocate metalinguistic discussions of

grammar (i.e., discussions using the grammar terms) but do draw attention to structural phenomena through focused practice on limited language samples, sometimes presented with the aid of special equipment such as cuisiniere rods and fidels. Alternatively, in the Suggestopedia approach, Lozanov (1982) while emphasizing the psychological component of language acquisition through a controlled environment and the use of a persona for the learner, does include conscious attention to aspects of language contrasts, at least in the early stages when long dialogues are presented to learners both in L2 and L1 translations. Doff (1988), in a training course for language teachers, advocates both contextualization and isolation of grammar elements in second language teaching with an emphasis on making grammar meaningful to learners. He suggests a limited role for the metalinguistic aspect of teaching through contrasting L2 grammar items with each other and occasional grammatical explanations in the first language. His justification for this is that the level of complexity in a grammar explanation may exceed the learner's ability to understand it via the second language.

Views Advocating a Role for Explicit Grammar Instruction

A number of methodologists and linguists come out squarely in favor of the advantages of conscious attention to grammar in L2 instruction, at least for some populations under specific conditions. It should be pointed out that these theorists do not claim that conscious grammar knowledge is a sufficient condition for learning and effective L2 use, only that it may be helpful and has the potential to accelerate L2 learning and acquisition and increase accuracy in production. Among these are Eisenstein (1987), Rutherford and Sharwood Smith (1988), Long (1988), and Celce-Murcia (1991).

Rutherford and Sharwood Smith (1988) suggest the "interface hypothesis." This claims that a conscious understanding of structural regularities in the language can affect the growth of subconscious grammatical knowledge in the learner's mind. Bialystok (1994) articulates a framework for explaining L2 acquisition and use which incorporates analysis of knowledge and control of processing. According to this view, "increasing analysis of mental representations ...leads to an increase in accessibility to knowledge" (p.160). This would contribute to an explanation of why conscious reflection on how language works could support second language acquisition and production.

Specifically, in applying her framework to second language instruction, Bialystok concludes that different kinds of language instruc-

1249

tion lead to concentrating on analysis, control, or both. Conscious grammar learning would develop analysis while meaning-focused practice would support meaning-based language processing and control.

To shed additional light on the effects of grammar teaching, some investigators have undertaken classroom-based research to study the role of conscious grammar teaching in the second language class. Montgomery and Eisenstein (1986) found that a course based on field trips and focused on language use with limited feedback on form promoted the development of grammatical accuracy for a group of adult learners who contemporaneously attended a more traditional ESL class. Learners who attended the same traditional class but did not participate in the field trips course made significantly less progress in the development of grammatical accuracy. This study was neutral, however, to the question of whether communicative input and opportunity for practice are sufficient conditions for SLA. Instead, this shows that the combination of metalinguistic language learning and a reality-based communicative environment can enhance interlanguage grammar development.

In addition to addressing the question of whether or not to incorporate conscious attention to L2 grammar at all, teachers must decide to what extent doing so may be helpful. Whether to present language awareness activities inductively or deductively, and how to alternate and integrate awareness activities with practice is a decision language teachers must grapple with constantly. A recent approach to teaching academic writing has been developed by Camhi (1993). Learners begin with a series of focused language awareness activities which incorporate a limited and internally consistent set of surface-level grammar rules. Once mastered conceptually, the metalinguistic terminology and rules are applied to a writing process approach in which learners are responsible for producing multiple drafts and keeping a personal grammar log to monitor problems and progress. Preliminary data indicate that for population samples from several urban colleges, pass rates on a holistically scored writing test are dramatically higher than for any other approach that has been tried to date.

Doughty (1991) investigated acquisition of restrictive relative clauses under conditions of meaning-oriented versus rule-oriented instruction. While both approaches resulted in gains, highlighting structures in context was more effective in promoting greater comprehension of input. VanPatten and Oikkenon (1996) in a study of the acquisition of Spanish pronouns and word order interpret their results to indicate that structured input was helpful while providing explicit information did not appear to be necessary.

1250

Do we agree about anything? Many of us believe that a knowledge of grammar rules cannot be a sufficient condition for language use; for some learners, pedagogical grammars may be aids to learning. In light of the lack of theoretical and empirical consensus reflected in our above discussion, it is crucial that we add to our understanding the voices of classroom teachers who face this problem on a daily basis and have developed working solutions for the populations they address. A reasonable place to start is to consider the experience of college-level teachers who are working with the community of learners most likely to benefit from a metalinguistic treatment of grammar.

A Study of Teachers' Views

Procedure

The authors of this article administered questionnaires to 60 ESL teachers of adult learners for academic purposes, 30 who were from the New York area and to the same number from Puerto Rico. The teachers had a range of experience from a minimum of three years to a maximum of 45 years. Teachers from New York had an average of 14.1 years of ESL experience; their counterparts from Puerto Rico had taught ESL for an average of 14.5 years. Of course, this is a limited sample intended as a first step in reflecting teachers' practice and viewpoints. As such, generalization to other communities of teachers and learners would not be appropriate.

Our questionnaire asked participants to respond to the following questions: *Should grammar be consciously taught in second language courses (yes, no, or sometimes)? If so, when and how? If not, why not? In either case, clearly explain your teaching approach and your rationale for it.* Teachers were also asked to rate their confidence in their approach on a three-point scale: (1) unsure, (2) experimental, and (3) well-developed. We also asked them to list any factors in their backgrounds which they thought might have influenced their approaches to grammar in language teaching. The definition of grammar was purposefully left to the respondent to interpret. To provide additional information and to triangulate data (Allwright & Bailey, 1991), open-ended (unstructured) post-hoc interviews (Bernard, 1995; Spradley, 1979) were conducted with eight teachers from each group who had taught for different lengths of time, in order to explore ideas and issues mentioned in the questionnaires more fully. Two experts in second language pedagogy, one from Puerto Rico and one from New York, were also consulted in order to add to our understanding of the data

obtained. Both experts possess doctorates in an area of applied linguistics and have over fifteen years of second or foreign language teaching experience.

Findings

In responding to what it meant to them, participants in the study defined grammar teaching in terms of syntactic elements which were essentially sentence level phenomena, with the exception of structures such as determiners which might be semantically linked to previously stated or assumed information. While experts noted that there is substantial research regarding discourse level and generative grammar, teachers did not refer to these views of grammar in the current discussion.

Regarding the general question of whether or not grammar should be taught, the responses of the 2 groups of teachers were distributed as follows:

	Yes	No	Sometimes
NY Teachers:	9	10	11
PR Teachers:	14	3	13

Comments on the questionnaires and post-hoc discussions with participants and language teaching experts in Puerto Rico and New York indicated general agreement that the apparently more positive and definite view on the part of some Puerto Rican as compared to New York based teachers regarding the role of conscious grammar was related to two major issues: the first is the more traditional approach to language teaching in general advocated in the pedagogical culture of Puerto Rico; the second relates to the fact that the Puerto Rico setting is more of an EFL than an ESL setting so that out-of-class English input would be more restricted. This might cause teachers to rely more on formal presentations including conscious grammar which might be acquired inductively in the mainland American second language setting.

> In Puerto Rico we respect tradition; we're careful about implementing new ideas just because they're new. Grammar has always been part of our language learning experience. We see no reason to abandon it totally. (Puerto Rican teacher, interview)

1252

> I believe that substantial access to input in the English speaking community explains the belief of many teachers and theorists in the U.S. that formal grammar presentation is unnecessary. (New York expert)

Regarding teachers' expressed levels of confidence with respect to their approaches to grammar teaching, the average degree of confidence in their approach expressed by each group was as follows:

	Unsure	Experimental	Defined Approach
NY teachers	4	10	16
PR teachers	1	4	25

Clearly, while teachers in Puerto Rico express somewhat more confidence in their approaches to grammar, a substantial number of New York teachers are also confident. Comments by respondents along with the explanations of our experts lead us to postulate that this high level of confidence comes from several sources. First, the issue of grammar teaching has high perceptual saliency for second language teachers. Several of our participants commented on their experience as language learners using conscious grammar insights in a positive way. In fact, all teachers we interviewed and both experts suggested that learning grammar consciously was a part of the educational experience of everyone in our study. Furthermore, several teachers noted that their students requested grammar instruction. In addition, in their experiences as teachers in training, the majority of participants noted that the issue of pedagogical grammar was an important part of their preparation and course work. This makes the nature of grammar teaching an issue that every teacher had to address from the beginning of their language teaching experience.

> When I was working on my Masters, we used to debate a lot about teaching grammar in our ESL classes. Everybody had studied some grammar. Personally, I enjoyed it. Some of us had a real love-hate feeling about it. Eventually, I guess we all have to make up our own minds about it. (New York teacher, questionnaire)

It is interesting to note that relatively more New York teachers rated their approaches to grammar as experimental and a few more characterized themselves as unsure, as compared with their counterparts in Puerto Rico. The primary explanation for this difference is

1253

attributed by participants and experts to two factors. First, the nature of college ESL classes in New York is that they contain substantial diversity in learners' language and cultural backgrounds. This means that approaches to grammar may differ in the degree to which they complement learners' previous experiences with how to learn languages. As a result, teachers are more likely to need an experimental view of approaches to teaching grammar to particular learners. This position coincides with Harley's (1993) view that grammar teaching must be flexibly geared to specific learners. Secondly, " teachers in the New York area are virtually inundated with opportunities to attend conferences and have contact with colleagues from around the world" (New York expert). For those teachers that take advantage of this setting, new ideas on language pedagogy in general and grammar in particular are constantly available for experimentation. In addition to the differential levels of confidence reported by our participants regarding their approaches to grammar teaching, there were distinct differences in opinion as to optimal classroom practice.

Three important questions in the discussion of the role of conscious instruction in second language teaching/learning which emerged from the data are:

* Should grammar be consciously taught and what should the relative balance between conscious instruction and communicative activities be?
* What should come first, conscious instruction or communicative activities?
* How should grammar be taught?

Here we present a qualitative analysis of the opinions teachers expressed with respect to these three questions integrating information from the written surveys and the interviews.

Interestingly, while there were differences among teachers regarding the degree to which grammar teaching in second or foreign language classes was appropriate, there were few clear distinctions between the language teachers in Puerto Rico and the continental US in the rationales expressed for particular views. The discussion below seeks to present the range of ideas and responses offered by our participants.

Should Grammar Be Taught?

Teachers' answers to the question of whether or not grammatical structure should be consciously taught ranged from "never" to

"always," with the majority of the respondents advocating some use of conscious instruction combined with communicative practice. The variables affecting the decision of how much conscious instruction should be used were the student's level of English proficiency, their age, their education level, and the frequency with which the items appear in language use, and whether or not a particular grammar problem represented a problem shared by many learners. With respect to student level, there was some disagreement. Seven teachers (three from Puerto Rico and four from New York) felt it was advisable to begin conscious instruction of grammar early on in the learners' career. "You can't start too early. If students learn correct grammar and practice it from the beginning, they will have a good foundation to build on" (P.R. teacher, interview). However, two teachers from New York and three from Puerto Rico felt that increasing and more complex grammar work was needed only as proficiency developed. Yet some (four teachers from Puerto Rico and one from New York) felt grammar should be taught at all levels. One New York teacher expressed the belief that different techniques for presenting grammar should be used at different levels.

As for age, despite a few dissenters, there was general consensus among the teachers and experts that conscious instruction should not be used with younger learners particularly from kindergarten to third grade. Many thought conscious grammar was appropriate for use with adults because they are highly motivated to know the rules and to monitor. More than one teacher expressed the opinion that a lack of teacher response to the perceived need for grammar can interfere with learning in adults. Another teacher agreed but added that conscious instruction is but one adult learning strategy among many and must be exploited. " My students are adults who often know the grammar of their native language and benefit from this kind of meta-cognitive information." (N.Y. teacher, interview) "My students want grammar; I believe they need it." (P.R. teacher, questionnaire).

When Should Grammar Be Taught?

As for when explicit grammar should be taught, one teacher from Puerto Rico and five from New York commented that they only teach grammar items which occur frequently in native speaker use. This, of course, is an impressionistic judgment the teacher must make. Others only focus on grammar problems which they have observed to be shared by a substantial number of their students. Conscious instruction is seen as a way of clearing up student doubts and as a way of building confi-

255

dence among learners.

> I usually try to teach English implicity. But if there is something unclear or confusing, I supply a mini-lesson: explicit grammar with rules and lots of examples...The mini lesson lowers the anxiety level. Often my students have been taught explicitly before, so some connection to what they are used to also seems to help. (PR teacher, questionnaire)

A good number of teachers only deal with grammar when it comes up naturally in the course of a communicative lesson. Grammar instruction is seen, however, as a way to create a shared body of knowledge which can be referred to later in language learning activities. The two experts both commented that the decision depends on the particular grammar feature and whether or not it lends itself to conscious teaching. The teaching of the negative was mentioned as an example of one such feature. Finally, we can say that there were some comments that whether or not grammar is presented consciously, people construct grammar themselves from data they get from a variety of sources.

> I would not consciously teach grammar, no. Students can easily get bogged down with too many rules and exceptions. I use thematic units in teaching. I feel that using a whole language approach exposes students to a broad range of language. (N.Y. teacher, interview)

Why Should Grammar Be Taught?

With respect to why grammar should be dealt with explicitly, several teachers (3 Puerto Ricans, 1 New Yorker, and both experts) felt that to be an educated person you need to learn grammar and that conscious instruction was appropriate among learners of a high educational level. Other factors which these teachers identified as affecting the decision about conscious instruction were student expectations and syllabus expectations. Many (5 Puerto Ricans and 8 New Yorkers) said they teach grammar because their students want or even demand it (although several teachers expressed the opinion that it was otherwise probably unnecessary). One teacher commented that her students thank her when she teaches grammar. Interestingly, a good number (9) of Puerto Rican teachers in particular mentioned syllabus expectations

256

as being the reason for which they taught grammar. One New York teacher mentioned these expectations as a factor in her decision making about conscious instruction for her Russian-born students. She felt that addressing the metalinguistic level of language was necessary in order for the students to "take her class seriously" (N.Y. teacher, questionnaire). Further, she was also among those who expressed the belief that students who have past experience learning languages with the support of conscious grammar rules will benefit from such an approach.

Another factor which was identified by teachers was the purpose for which the student was learning English. If the students' goal is to use English academically or in the work place, it was felt that greater emphasis on grammar instruction was appropriate. "I teach some elements of grammar explicitly to adult and university-level students and/or those bound for business or professional careers. I believe this is appropriate for their needs" (P.R. teacher, questionnaire).

Both experts concurred that in these contexts clarity and correction are both factors which can influence the level of success attained by the learner. " It would seem that many practitioners have come to believe that individuals whose goal is to develop excellent English for use in academic or business environments can achieve greater accuracy and control through some measure of grammar instruction" (N.Y. expert).

It was also felt that explicit knowledge of grammar is especially useful to monitor one's formal writing. Thus, conscious instruction of grammar is seen as a way of improving students' use of formal English. "I give them what they, in my opinion, must have in order to continue to make progress in formal English use, in writing especially and in language's other parts as well" (N.Y. teacher, questionnaire).

Finally, we will note that two teachers said that they teach grammar because the conscious instruction of grammar worked for them as individuals when they learned a second language, and they trusted their own experience more than the theories of researchers. "My own education included very formal language study including memorization, reading, writing , and grammar. Now I 'm using a communicative approach, but I won't completely abandon the teaching that worked for me. Grammar helped me and I can see that it also helps my students. I have confidence in my own experience" (PR teacher, interview).

What Should Be Dealt with First?

As for our second question of what should be dealt with first, conscious instruction or communication in meaningful contexts, this proved

1257

to be controversial among experienced teachers. Teachers were quite divided in their opinions whether to begin with instruction or communication. Some (6 Puerto Ricans and 11 New Yorkers) felt that in the students' learning careers, it was better to begin with communication and to introduce conscious instruction as a corrective later on. Others (5 Puerto Ricans and 2 New Yorkers) felt just the opposite, that conscious instruction should come first. Alternatively, several teachers advocated a spiraling or recursive relationship between conscious instruction and communicative activities.

Thus, teachers came to no consensus with respect to which should come first, consciousness raising or communication. Yet an important conclusion was that most felt that both have a role, irrespective of the order of presentation. A judicious blend of both activity types was seen as the best way to foster second language learning and acquisition. "I try to integrate grammar while emphasizing meaning. I compromise between giving enough practice so it's internalized but alternate with explanations when needed, especially when meaning hinges on it" (NY teacher questionnaire).

There was also disagreement about whether conscious grammar teaching was more appropriate in early versus later stages of second language proficiency. This controversy mirrors R. Ellis' (1994) discussion of the continuing debate between researchers such Long and Krashen as to whether beginners benefit more than advanced learners from conscious attention to grammar, or whether the reverse may be true (pp 615-617). Like the researchers, teachers continue to explore this issue.

The teachers also differed in their assumptions about the degree to which individuals pass through predictable developmental stages of grammar acquisition as has been postulated by some theoreticians and researchers. Among those who expressed a belief in predictable stages there was the sense that a grammatical syllabus should mirror this "natural order of acquisition." One teacher stated an opposing view. " I feel that conscious grammar analysis and practice are especially appropriate for grammar items which tend to be acquired late but for which learners might be penalized" (N.Y. teacher, interview).

How Should Grammar Be Taught?

What of the teachers' opinions about our third question, the "how" in grammar teaching? Of those who commented on this issue, most (7 Puerto Ricans and 12 New Yorkers) teachers felt that any focus on grammar should be contextualized. One teacher said that she only

teaches the syntactic dimension and leaves it to students to make the interface with the semantic, rhetorical, and phonological components through contextualized use. Some respondents (4 Puerto Ricans and 2 New Yorkers) specifically recommended working with grammar in conjunction with reading. Writing was also seen as an important context in which conscious instruction can be exploited, particularly in the editing phase of the writing process. Writing conferences were recommended as a technique for helping students inductively discover reasons for grammar choices. One teacher highlights grammar errors after students have produced their third draft of a piece of writing. Also, class discussion of grammar points was seen as a useful technique. One teacher suggested the use of humorous phrases or contexts as a way of presenting grammar points. A few felt (4 Puerto Ricans and 1 New Yorker) that contrastive analysis with the L1 was a valuable approach to helping students understand L2 grammar. "This may seem old fashioned, but I still use focused comparison of English with Spanish. This seems to clear up some difficult language problems for my students" (PR teacher, interview). Interestingly, one teacher also recommended the use of translation as a way of introducing conscious instruction of grammar points.

Collaborative learning was viewed as a useful tool in developing grammar knowledge. Several teachers and both experts recommended that students form peer teaching groups to work on revision and correction of writing. "In order to give language learners the opportunity to review and revise their work, peer feedback is a must. Individual conferences with the teacher can't do the whole job. Group work helps students become more aware of meaning, style, and grammatical accuracy in their writing" (NY Expert). Other suggestions included the collaborative learning of grammar through reference to a text which students would use to complete grammar tasks.

Ellis (1995) identifies an important distinction between the presentation of grammar elements to second language classes as an inherent part of the curriculum versus attention to the grammatical needs of specific learners as they arise. A number (3 Puerto Ricans and 7 New Yorkers) of the teachers surveyed incorporated the latter alternative. Camhi (1993) has each learner in ESL writing classes keep a log of his or her grammar problems which have been identified in writing samples along with the correct form and relevant comments. Having students keep their own grammar notebooks was also suggested by several of the teachers surveyed as a valuable approach along with the use of a grammar text for reference purposes when students need to consult it regarding a specific item or issue they find problematic.

Conclusion

In general we can conclude that the views of the teachers surveyed regarding conscious grammar teaching reflected the range of ideas and findings reported in the literature and advocated by methodologists and curriculum developers. Nevertheless, among Puerto Rican teachers there was little conflict regarding the need for some level of formal attention to grammar in the ESL classroom. This is consistent with the views of methodologists like Rutherford and Sharwood Smith (1988) and Fotos (1991). Although some New Yorker teachers shared this view, by contrast, a substantial number (10) advocated a purely meaning-focused approach to language teaching, believing some combination of comprehensible input and practice to be sufficient conditions for second language development. This view corresponds to the recommendations of Krashen (1996) and Krashen and Terrell (1983). As noted above, this distinction could conceivably be attributed to a difference in the educational cultures of the two communities, differing opportunities for learner input, or both.

We have found that teachers' ideas regarding grammar teaching were generally well developed. This was confirmed by their abilities to express clear and coherent rationales for their approaches to curriculum and pedagogy. Nevertheless, more of the New York teachers were unsure or experimental than was the case with the Puerto Rican teachers.

Reasons given for how and why conscious grammar was taught were based mostly on teachers' perceptions of their own experience as teachers and learners. It is interesting that our participants rarely justified their approaches by referring to research studies or any particular methodology. Nevertheless, most teachers thoughtfully reflected on what they believed was effective for their students and seemed to be making principled choices.

Given the lack of consensus among researchers, curriculum developers, and teachers, we will not attempt to recommend any particular pedagogical approaches at this time. Perhaps, given the complex range of variables involved in second language acquisition, the presence or absence of explicit grammar instruction while potentially useful in the right circumstances for the appropriate learners, is not crucial for acquisition. Alternatively, it may be the case that even in the most meaning-centered classes, some aspect of attention to grammatical form cannot be avoided. These are only two of the many possibilities that would provide an explanation for teachers' views.

What do the authors think? We must acknowledge that based on our own experience as formally educated multilingual learners and university teachers, our best guess at this point is that the reflective incorporation of grammar rules in a communicative class where there is substantial opportunity for interaction and feedback appears to be helpful for promoting comprehension and accuracy. However, in our view the matter is far from settled.

The natural next step has been suggested by numerous researchers and methodologists (Williams, 1995), namely the continued collaboration of teachers and researchers, as well as teachers as classroom researchers in the investigation of when and how conscious attention and metalinguistic consideration of grammar can promote and enhance learners' second language development. This information derived from classroom-based research will enrich approaches to teaching and will inform and validate our theoretical understanding of this complex and fascinating area of linguistics.

References

Allwright, D., & Bailey, K. (1991). *Focus on the classroom*. New York: Cambridge University Press.

Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Batstone, R. (1994). *Grammar*. Oxford: Oxford University Press

Bernard, H. R. (1995). *Research methods in anthropology*. Walnut Creek, CA: Altamira Press.

Bialystok, E. (1979). Implicit and explicit judgments of L2 grammaticality. *Language Learning, 29*, 81-103.

Bialystok, E. (1990). *Communication strategies: A psychological analysis of second language use*. Oxford, UK: Basil Blackwell.

Bialystok, E. (1994). Analysis and control of the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157-168..

Camhi, P. (1993). A metalinguistic grammatical approach to second language writing development. Presentation at Second Language Acquisition Circle. New York University, New York.

Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly, 15*(3), 459-480.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Doff, A. (1988). *Teach english: A training course for teachers*. Cambridge: Cambridge University Press.

1261

Doughty, K. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 12*(4), 431-470.

Eisenstein, M. (1987). Grammatical explanations in ESL: Teach the student, not the method. In Long, M. & Richards, J. (Eds.), *Methodology in TESOL* (pp. 282-297). New York: Newbury House/ Harper & Row. pp. 282—297.

Ellis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly,* (28) 87-107.

Ellis, N. (1994). *Implicit and explicit learning of language*. London: Academic Press.

Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.

Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness raising. *TESOL Quarterly*, *28*(2), 323-351.

Fotos, S. & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly, 25*(4), 606-628.

Gattegno, C. (1976). *The common sense of teaching foreign languages.* New York: Educational Solutions.

Harley, B. (1993). Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition, 15*, 245-259.

Howatt, A. P. R. (1984). *A history of English language teaching*. Oxford: Oxford University Press.

Kelly, L.G. (1969). *Twenty five centuries of language teaching*. Rowley, MA: Newbury House.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamen.

Krashen, S. (1985). *The input hypothesis: Issues and implications.* New York: Longman.

Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London: Academic Press.

Krashen, S. (1996). *Under attack: The case against bilingual education*. Culver City, CA: Language Education. Associates.

Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamino.

Lightbown, P., & Spada, N. (1990). Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, *12*, 429-446.

1262

Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge: Cambridge University Press.

Long, M. (1988). Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, R. Ginsberg, & C. Dramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp.39-52). Amsterdam: John Benjamins.

Long, M. & Porter, P. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, *19*(2), 207-227.

Long, M. & Crookes, C. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, *26*(1), 27-56.

Lozanov, G. (1982). Suggestology and suggestopedia. In Blair, R. W. (Ed.), *Innovative approaches to language teaching* (pp. 146-159). New York: Newbury House.

McLaughlin, B. (1990). Conscious vs. unconscious learning. *TESOL Quarterly. 24.* 617-634.

Montgomery, C., & Eisenstein, M. (1986). Real reality revisited: An experimental communicative course in ESL. *TESOL Quarterly, 19*(2), 317-333.

Nunan, D. (1989). *Understanding language classrooms: A guide for teacher-initiated action*. New York: Prentice Hall.

Nunan, D. (1990).Action research in the language classroom. In J. Richards & D. Nunan (Eds.), *Second language teacher education* (pp. 62-81)

Pennington, M. C. (1995). *New ways in teaching grammar*. Alexandria, VA: TESOL.

Pica, T. (1987). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning, 37*(4), 563-593.

Pica, T. (1994). Questions from the language classroom: Research perspectives. *TESOL Quarterly, 28(1),* 49-79.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition. 6*, 186-214.

Pienemann, M. (1987). Psychological constraints on the teachability of languages. In C. Pfaff, (Ed.), *First and second language acquisition processes* (pp.143-168). Rowley, MA: Newbury House.

Richards, J., & Rodgers, T. (1986). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning, 47*(1), 45-99.

Rutherford, W., & Sharwood Smith, M. (1988). *Grammar and second language teaching: A book of readings*. New York: Newbury/

1983

Harper & Row.

Savignon, S. J. (1991). Communicative language teaching: State of the art. *TESOL Quarterly, 25(2),* 261-277.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11(2),* 209-231.

Schulz, R. (1991). Second language acquisition theories and teaching strategies: How do they fit? *The Modern Language Journal, 75*(1), 17-26.

Sharwood Smith, M. (1993) Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15(2).* 165-179.

Spradley, J. P. (1979). *The ethnographic interview.* New York: Holt, Rinehart & Winston.

Richards, J., & Rodgers, T. (1986). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Stern, H. H. (1983). *Fundamental concepts of language teaching.* Oxford: Oxford University Press.

Trampe P. (1994). Rules, consciousness, and learning. In R. Barasch, & C. James (Eds.), *Beyond the monitor model: Comments on current theory and practice in second language acquisition* (pp.159-168). Boston: Heinle & Heinle.

VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*(20), 225-248.

VanPatten, B., & Oikkenon, 5. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition, 18*(4), 495-510.

Wenden, A. (1991). *Learner strategies for learner autonomy.* New York: Prentice Hall.

Williams, J. (1995). Focus on form in communicative language teaching: Research findings and the classroom teacher. *TESOL Journal, 4*(4), 12-16.

Woods, D. (1996). *Teacher cognition in language teaching.* Cambridge: Cambridge University Press.

1264

Authors

MIRIAM EISENSTEIN EBSWORTH, Director of Doctoral Programs in Multilingual Multicultural Studies, Department of Teaching and Learning, New York University, New York. E-mail: Eisenstn@is3.nyu.edu. *Specializations*: sociolinguistics, second language acquisition.

C. WILLIAM SCHWEERS, Associate Professor of English as a Second Language. University of Puerto Rico, Bayamon, Puerto Rico.E-mail: CW_UPR.CUTB.CLU.EOU. *Specializations*: second language policy, ESL.

1265

Applied Language Learning
1997, Vol. 8, No. 2, pp. 261-284

Proficiency Guidelines and Language Curriculum
Making ACTFL Proficiency Guidelines Effective in Furthering Japanese Language Proficiency[1]

Yuki Johnson
University of Michigan

This article examines the role of ACTFL Proficiency Guidelines as a system of language proficiency testing and discusses how effective that system is in relation to classroom instruction at the university level. There are two main issues addressed in regard to adapting guidelines into language curriculum.

1. Viewing ACTFL Proficiency Tests as having the characteristics of achievement tests. Once the ACTFL guidelines are adapted into a specific language curriculum, this problem must be addressed since the guidelines cease to function as proficiency testing. In order to reap the greatest possible results from the relationship between proficiency tests and achievement tests, curriculum should be modified so as to embrace a proficiency-oriented instruction which aims to cultivate learners' communicative competence. The type of testing that would result I refer to by the amalgam "achievement-proficiency" test.

2. The ACTFL Guideline's rating scale. Major components of communicative competence are not properly reflected in the scale even though the goal of testing is to measure this competence. An alternative version of the rating scale is proposed that incorporates theoretical frameworks suggested by Bachman and others and that builds on a strong relationship to classroom instruction. When the various theoretical and practical issues introduced here are taken into consideration, the ACTFL guidelines, in their unavoidable guise as achievement-proficiency tests, should result in greater reliability and validity than general proficiency tests and should serve as a helpful tool to aid in the process of language acquisition.

© *1997, Yuki Johnson*

1266

Yuki Johnson

The Role of ACTFL Proficiency Guidelines in Language Proficiency Testing

The history of the proficiency testing movement can be traced from its foundation in one government agency (the Foreign Service Institute) (FSI), through its development with cooperating government agencies (the Interagency Language Roundtable) (ILR), to its expansion into academia through the work of the American Council on the Teaching of Foreign Languages (ACTFL) and the Educational Testing Service (ETS) (Lowe & Stansfield, 1988).

The resulting set of ACTFL Proficiency Guidelines are primarily skill-level descriptions aimed at college students. They were produced specifically to help language educators measure learners' functional abilities (creative language use in untailored situations) in a target language instead of merely assessing what learners know about that language (knowledge of phonology and structural devices of the language, for example). The ACTFL organization presents its proficiency guidelines as:

> a hierarchy of global characterizations of integrated performance in speaking, listening, reading, and writing. This is not equal to an achievement test which assesses a learner's attainment of various specific aspects of course and curriculum content. Each description is a representative, not an exhaustive, sample of a particular range of ability, and each level assumes all previous levels, moving from simple to complex in an "all before-and-more" fashion.

As a result, the guidelines are designed to assess:

> functional language skills through a criterion-referenced, rather than a norm-referenced assessment: the global tasks or functions performed with the language; the social contexts and the content areas in which the language can be used; the accuracy features which define how well the learner performs the tasks pertinent to those contexts and content areas. Each testing sample is rated solely according to the criteria of the rating scale. (ACTFL Proficiency Guidelines, 1986)

In the last decade, the ACTFL guidelines have become the prevailing road map for language testing at American and Japanese colleges and universities. With the increased awareness of this method by

language educators comes the promise of relatively uniform testing procedures and thus the possibility of a globally integrated system.

There are certainly extensive benefits in adhering to a set of language proficiency guidelines such as those presented by ACTFL. For instance, many institutions struggle with placing new or transferred students into the appropriate level of study for a particular language. Depending on past textbooks or teaching methods, students come to a school with divergent backgrounds relative to the acquisition of their target language. Students' courses of study may have approached grammar items in a different order than that followed by their new institution, for example, and the vocabulary and expressions they acquired may not mesh well with a given school's course levels.

When this problem occurs, the guidelines can help instructors discern which ACTFL level the learner has attained. Theoretically, one could use the guidelines to assess learners' general language competence regardless of background variables such as learning experiences and the use of various textbooks. This measure also lets instructors communicate more accurately with other instructors concerning new learners.

The guidelines assist with more than just placement, however, as there are more direct benefits to learners as well. The guidelines offer feedback that can help in advising learners how their course of study is progressing, how much their language skills have improved, what areas remain undeveloped, and so forth. When there are specific criteria in place to inform learners where they stand in relation to a particular knowledge base, these criteria can act as encouragement. In a perfect world, the guidelines would provide a measuring stick for both evaluators and learners.

Problems arise, however, if the ACTFL Guidelines are adopted in an individual language course curriculum to measure students' achievement of the course of study. Educators end up either modifying and adjusting the Guidelines to their curriculum or modifying the entire curriculum and teaching approach to the Guidelines. Either way, the Guidelines' original characteristics and the intentions of those who wrote them are altered. I witnessed the latter at the University of Minnesota, and the rest of the article addresses the problems that are seen when such a theory is put into practice.

The Relationship Between ACTFL Proficiency Guidelines and Language Instruction

Beginning in 1993, University of Minnesota students were required

291268

to take a proficiency test following their sixth quarter of language study if that language study were to be used to fulfill the University's College of Liberal Arts' language requirement[2]. A student would need the class credits to fulfill the language requirement, but also need to pass the proficiency test in order to graduate. When such a test is tied to students' graduation requirements, the relationship between the course curriculum and the proficiency test cannot be ignored.

Once an entire university language program adopts the ACTFL Proficiency Guidelines, its curriculum will need to be modified to facilitate sufficient performance on such a test. Since the proficiency test described above is decisive in terms of who does and who does not meet graduation requirements, it is natural for learners to expect that instructors relate course syllabi to the final proficiency test. In order to properly prepare students under such circumstances, course work must consistently aim toward a goal of proficiency.

Being proficient in a language means that a person can use the language accurately to communicate — i.e., the person has communicative competence.[3] This ability must include the phonology and structural devices of the language, as well as discourse, sociolinguistic, and strategic competence, to be considered complete mastery of all the components of communication. If the goal of a course is primarily to attain only one component of the various language skills (reading, writing, listening, and speaking) through such means as translating sentences or memorizing dialogues in a text, for example, then a proficiency test does not motivate learners.

A proficiency test would also lack validation, since the concept of proficiency does not refer merely to the acquisition of translation or memorization skills, but rather to acquiring overall language skills. To facilitate mastery of these skills, the learning environment should be one which cultivates such learner communicative competence. To achieve the best possible results, I would like to suggest that the communicative approach, as described below, is the teaching approach most compatible with proficiency-oriented instruction.

Proficiency-Oriented Instruction and The Communicative Approach

The word "communicative" was first used in its present respect by Campbell and Hymes (1972) to critique Chomsky's "Competence and Performance Dichotomy" in second language acquisition. As they initially pointed out, and as is clarified here by Omaggio (1993), Chomsky's dichotomy does not include any reference to either the

1269

appropriateness of an utterance to a particular situation or context or its sociocultural significance (p. 3). To fill this void, Campbell and Wales (1970) and Hymes (1972) proposed a much broader view of competence, one in which grammatical competence is merely one component of the overall knowledge a native speaker possesses. They termed this broader notion communicative competence and incorporated into that concept the knowledge of the rules of language use, including sociolinguistic and contextual competence in addition to grammatical competence. They also maintained the theoretical distinction between communicative competence and communicative performance, the latter being a term that reflected language in actual use, with all of the performance errors that derive from nonlinguistic factors.

In the time since Campbell and Wales' and Hymes' proposals, these descriptions have come under debate among many educators, sociolinguists, philosophers, and others. Of course there have been subsequent changes in the generally accepted definition of communicative competence.

In 1976, the language educator D. A. Wilkin's "Notional Syllabuses" had a significant impact on the development of a communicative approach to communicative competence, an approach which he labeled "communicative language teaching." The rapid application of his ideas by textbook writers, and equally rapid acceptance of these new principles by British and American language teaching specialists, curriculum development centers, and even governments, gave prominence nationally and internationally to what came to be referred to as the communicative approach.

There is no single text or authority on it, nor any single model that is universally accepted as authoritative (Richards & Rogers, 1986). Often set up as opposing the Audio Lingual Method, it is not a rigidly prescribed set of methods, but exists as a flexible teaching process. It is in fact relatively difficult to define what the communicative approach is systematically. The definition, content, and explanation of it differs from person to person. Nevertheless, there is some common ground for us to tread.

Owing to the work of Richards and Rogers (1986), it is possible to discern three core elements in the communicative approach to teaching: (1) activities that involve real communication promote learning, (2) activities in which language is used for carrying out meaningful tasks promote learning (Johnson, 1982), and (3) language that is meaningful to the learner supports the learning process (Richards & Rogers, 1986, p. 72). It is these factors that detail the spaces in which the

CAS 1270

communicative approach is capable of being compatible with proficiency-oriented instruction.

Proficiency-oriented instruction is the over-arching, goal-oriented (as, for example, is seen in the ACTFL rating scale) blueprint that a language program can be designed around. The communicative approach is a process-oriented (used as an approach in the daily classroom) teaching approach that puts that blueprint into practice. Both proficiency-oriented instruction, as a guide, and the communicative approach, as a practice, aim for accurate, creative language use in real-life language situations (authenticity). Learning activities are consequently selected according to how well they engage the learner in meaningful and authentic language use. Consequently, the communicative approach is used to improve learners' real-life performance — that is the way they perform in authentic language use situations.

One thing to keep in mind is that since the communicative approach is not a circumscribing *method*, but is rather a relatively loose-knit *approach*, instructors often seem to find enough slack in it to entangle both themselves and their learners.

The most significant misunderstanding is the idea that grammar explanations and the correcting of pronunciation should not be attempted during a class period. Taken this way, the communicative approach can cause problems: the idea of task-oriented class periods is pushed far to one side and the focus is put on how to manage tasks and games in class rather than what to teach.

Also, instructors seek the best ways to ensure students have fun in class, but not the best ways to answer their questions on linguistic matters. Although some linguists are still debating whether grammatical competence should be considered one of the components of communicative competence, it is now understood that the communicative approach provides a more viable example of the cognitive skills required to use language than did the earlier skills and components models, for example. This is because the communicative approach emphasizes language as a dynamic process, the way in which language is used to satisfy communication goals, and adherence to knowledge of grammar rules. Following the communicative approach in the classroom, instructors should take advantage of a variety of teaching methods. Choices should depend not only on what is taught, but also on how learners can best access the material.

Now that proficiency-oriented instruction and the communicative approach have been defined and their relationship to one another explored, the assessment problems that result from this relationship can be addressed. Taken together, proficiency-oriented instruction and the

5755271

communicative approach create a blurred vision of proficiency and achievement as a testing concept since, in this environment, proficiency tests and achievement tests would theoretically assess the same aspects of language use.

ACTFL Proficiency Tests as Achievement Tests

Achievement tests are often called attainment tests or mastery tests and are used to measure learners' attainment of particular course content. Instructors use them to confirm the amount and degree of control a learner has over materials presented within a specific period of instruction in a specific course. Proficiency tests, on the other hand, are used to measure a learner's control of a wide range of communication skills, such as use of grammar and vocabulary, phonological ability, and cultural insights, all irrespective of formal training. Bachman (1990) explains achievement tests as *syllabus-based* tests and proficiency tests as *theory based*-tests. The ACTFL Proficiency Guidelines state that "proficiency tests are not intended to measure what an individual has achieved through specific classroom instruction, but rather allow assessment of what an individual can and cannot do regardless of where, when, or how the language has been learned or acquired" (ACTFL Proficiency Guidelines, 1986).

As was mentioned earlier, in a case like that of the University of Minnesota, students have to take proficiency tests after attending a certain number of language courses. As these tests can prevent them from graduating, the course curriculum and syllabus must take that fact into account in order for proficiency tests to be truly valid. This is where it is important to note that the communicative approach and proficiency testing are backed by the same basic concept: communicative competence. Despite the theoretical distinctions between achievement tests and proficiency tests, practically speaking, when achievement tests in a particular course measure learners' communicative competence, they function as proficiency tests. In other words, achievement tests, which by definition assess curriculum-based performance after a certain period of language instruction, will also measure how proficient learners become in a classroom geared toward communicative competence. Nevertheless, educators need to observe the characteristics of these tests carefully and from the viewpoint of both tester and examinee.

Achievement tests are usually given to learners by the instructors who have taught those learners and are usually given after a certain predetermined amount of course instruction. The instructors have spe-

ERIC 272

cific expectations as to what the learners are supposed to be able to do, based on daily classroom activities. The instructors know the topics and the amount of vocabulary, grammatical rules, expressions, and cultural data that have been introduced to the learners who are being tested. Therefore, the content of their achievement tests is constructed from what the learners have already studied and is arguably equivalent for all peers in a given institution. Furthermore, learners can study for achievement tests by reviewing material presented during the particular period of instruction. They are aware that their peers have learned from the same materials and can, to some degree, predict what will be on a test. In this sense, achievement tests are norm-referenced. The results can be used for the preparation of instructional materials and reinforcement of areas of weakness.

Achievement tests, in this case, have a very basic function: regardless of the method or approach used in the course, instructors measure learners' retention of the specific materials introduced in class. Instructors may measure how many Chinese characters learners have memorized, for example, or how well they have retained knowledge of specific grammatical structures. This kind of achievement test results in what Rea (1985) originally called "non-communicative performance." Non-communicative performance is, however, best defined by Brindley who noted:

> it [non-communicative performance] is not focused on real-life communication but on the enabling skills that allow communication to take place. Instructor-made tests such as dictation, clause, or vocabulary tests might enable instructors to find out about a learner's mastery or nonmastery of particular course-content, but would not generally give a direct indication of ability to mobilize these enabling skills for communicative purposes. (1989, p. 12)

Proficiency tests do not possess these characteristics. First, as ACTFL Proficiency Guidelines suggest, proficiency tests should be given by testers who do not have any background information on the learners being tested. If they know the learners, the content of the tests tends to be biased (for example, a situation in which an instructor may not ask something they are sure from experience a particular learner would not know). Theoretically, textbooks, location, and the length of learning periods are outside of consideration in proficiency testing, while achievement tests are made based on these elements. When, for example, one conducts an oral proficiency interview (OPI) with some-

1273

one one knows, it is almost impossible to avoid considering the past learner-instructor relationship. If an instructor asks something s/he is confident the examinee does not know, the instructor can feel like s/he is being intentionally cruel. If an instructor explores only things the examinee knows, s/he voids the purpose of the examination. Such an interview lacks reliability. It should also be noted that even if learners are told they will take a proficiency test, they do not know what they are going to be asked — i.e. the content of the test is relatively unpredictable. This can be a heavy psychological burden and can lead examinees to feel uncomfortable, rendering some unable to perform at their "real" ability level. (Others may in fact thrive under this burden.)

Although there are a number of factors that differentiate achievement tests from proficiency tests, when tied to a course of instruction meant to cultivate learners' communicative competence (proficiency-oriented instruction), these tests overlap in that both measure what learners can do using language. In the communicative approach to reading, for example, learners are taught how to skim through authentic reading materials such as advertisements, magazines, and newspaper articles, as well as how to do more intensive reading. In writing, students learn how to form and use meaningful sentences for writing short memos, thank-you cards, letters of intent, theoretical papers, and so on. These materials relate to authentic language use in the classroom as they are "real-life" materials. Learners are tested on what they have learned in class, and therefore on how proficient they have become in the authentic use of language. I would suggest that this type of achievement testing coupled with this type of curriculum unavoidably creates a different type of proficiency test in a narrow sense — a hybrid I refer to as an "Achievement Proficiency Test."

Viewed from the other direction, it is equally easy to see how a proficiency test fits the broad definition of an achievement test when instructors use it as a yardstick to measure learners' retention of specific materials introduced in class. When used in conjunction with a program that adheres to communicative competence in its classrooms, the ACTFL Proficiency Test becomes in reality this same Achievement-Proficiency Test.

The following is a concise description of the test types mentioned.

1274

Achievement	Measures achievement of specific course objectives. Relates only to material presented in the course by the instructor.
Proficiency	Measures global proficiency: what the learner can do using the language, irrespective of the learners' learning experiences.
Achievement-Proficiency	Measures achievement of one's proficiency in language use. Relates to authentic real-life materials which are presented based on an approach that aims to cultivate learners' communicative competence.

Figure 1

If an educational institution adopts a specific teaching method, such as the Grammar Translation Method or the Audiolingual Method — i.e., one that focuses only on specific aspects of language use, such as language structure and habit formation — neither Proficiency Tests nor Achievement-Proficiency Tests are valid. Singularly focused approaches go hand-in-hand with achievement tests.

If an institution, of any type, wants to measure an examinee's target-language ability, it properly adopts a proficiency test. An examinee may work in the stock market, be a college student, a biologist, and so on. Learners' experiences vary and therefore the items they know or can use also vary. The tester tries to measure how proficient an examinee is in the language, regardless of his/her background.

An educational institution that has adopted the ACTFL Proficiency Guidelines is automatically involving itself in assessments that are hybrids of achievement and proficiency tests. The aim is to measure communicative competence gained primarily through the material presented in its own students' course of study. Once an institution adopts the ACTFL Proficiency Guidelines into its curriculum, it is inevitable that the ACTFL Proficiency Tests have a trace of achievement test in them, although that may not be what ACTFL intended.

Authenticity in Achievement-Proficiency Tests

Proceeding with the notion that in proficiency-oriented instruction curriculum courses will best be taught using the communicative approach, and learners' progress will be assessed using an achievement-

1275

proficiency test, the specific content of such a test in this learning environment needs to be explored. ACTFL states that ACTFL Proficiency Guidelines assess "the ability to use language effectively and appropriately in real-life situations." The guidelines are intended to assess communicative competence. This means that the content of the ACTFL tests needs to focus on authenticity, a requirement that extends to achievement-proficiency tests. Bachman (1990) suggests that authenticity and proficiency are synonymous when a teaching approach attempts to duplicate real-life situations in which proficiency is normally demonstrated (p. 303). This is seen in the communicative approach which aims at cultivating communicative competence (including the often ignored subcomponent, grammatical competence.)

Consequently, classroom instruction must approach communicative competence in not only top-down, but also bottom-up approaches.[4] Further, because (as was previously explained) the achievement-proficiency tests are inextricably linked to graduation requirements, they must reflect course curriculum. This means that achievement-proficiency tests may also need to include not-necessarily real-life factors such as cloze, dictation and recall protocol, for example. The exclusive use of real-life situations in testing makes it difficult to define the domain of non-test performance as distinguished from test performance. This is better illustrated by a Cohen testing diagram. (1994, p. 162)

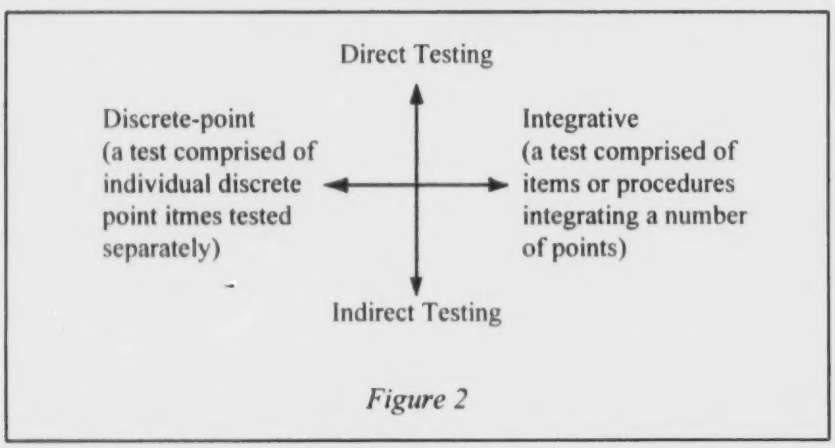


Figure 2

As per the concepts thus far presented, an ACTFL OPI would lie in the quadrant between direct testing and integrative testing, both of which assess a learner's global speaking ability. Cloze testing would fall in between integrative and indirect testing; traditional multiple-choice grammar tests would fall in between discrete point and indirect point testing; and testing for verb conjugation in speaking would fall

in between direct testing and discrete point tests. As all four areas noted on the diagram should be emphasized equally in achievement-proficiency tests, achievement-proficiency testing is located at the point where all four components intersect on the diagram.

The limitations of the autonomous use of either "not necessarily real-life factors" or "real-life factors" becomes more apparent in reviewing two issues Bachman (1990) presents in relation to the real-life approach:

1. The failure to distinguish between the trait or ability measured, and the observation of behavior limits the interpretation and use of test results.
2. A real-life approach provides an inadequate basis for examining validity.

This can also be taken as evidence that authenticity is related to the concept of achievement. The point here is, where and how do learners acquire such authenticity? They predominantly acquire it through classroom materials and activities, though this may not necessarily be the case if they are in a country where the target language is spoken.

Suppose that a test includes only very authentic question items such as listening to recorded announcements from a Japanese department store, a train station, and so on. If learners have never experienced those types of environments, then it is very difficult for them to process such samples. In other words, real-life situations like these do not necessarily always assess real communicative competence, nor can they be used as evidence that examinees do not have a particular listening ability.

The same can be said for speaking, reading, and writing. When writing someone a letter in Japanese, one writes the name of the person the letter is addressed to, a phrase equivalent to "sincerely yours," and the date at the bottom of the letter. Learners cannot be expected to intuit differences like these between the way a letter is written in English and in Japanese. Obviously they have to learn from someone, or they have to be exposed to the culture and learn it informally through experience. Even if they are in Japan for a long period, however, it is not necessarily true they would learn how to write a letter. Also, in conversation, learners cannot know that Japanese people give responses (*aizuchi*) to almost every sentence they hear as a way of "smoothing out" their conversations.

In point of fact, there are a number of barriers to gaining such discourse (textual) competence, especially when learners are studying

in the United States and are surrounded by American culture. Authenticity in these cases will most likely be acquired through materials introduced in class. When tests include what is taught in class, and what is taught in class includes authentic materials, then both teaching and testing are geared toward authenticity.

Problems Associated with ACTFL Assessment Criteria in Japanese Language Achievement-Proficiency Tests

The practical aspects of the ACTFL Proficiency Guidelines (which in this article have been tied to course curriculum and are considered to be achievement-proficiency tests) are most well known in their manifestation as oral proficiency interviews (OPI), and primarily work to test speaking ability. The ACTFL rating scale used to assess speaking is based on a number of components of language performance (and, in one instance, subcomponents). The components of assessment include (ACTFL Proficiency Guidelines, 1989):

1. Global Tasks or Functions (refers to what speakers are able to do with the language).
2. Context (refers to circumstances or settings in which a person uses language).
3. Content (refers to topics or themes of conversation).
4. Accuracy (refers to the acceptability, quality, and precision of the message conveyed). (Accuracy is further subdivided to include: fluency, grammar, pragmatic competence, pronunciation, sociolinguistic competence, and vocabulary).
5. Text Type (refers to the quantity and the organizational aspects of speech produced by the interviewee).

Note the following diagram.

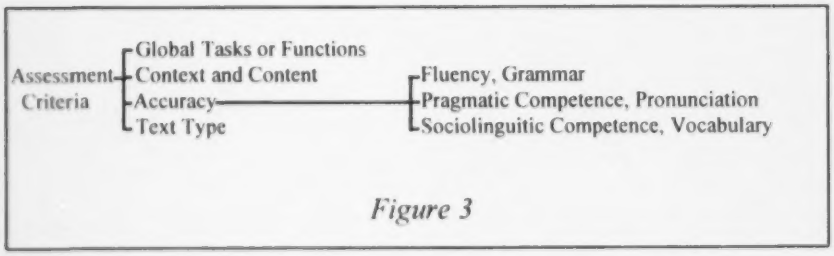


Figure 3

NSU275

The problem with these assessment criteria is that fluency, grammar, pragmatic competence, pronunciation, sociolinguistic competence, and vocabulary are too dissimilar of criteria to be lumped together under the rubric "accuracy." Grammar, vocabulary, and pronunciation are basic constituents of language use and as such can be assessed in a relatively straightforward manner. Sociolinguistic competence, pragmatic competence, and fluency are much more subtle and complex criteria which therefore require a more thorough and more detailed assessment[5].

Interested Japanese language educators participating in ACTFL workshops developed Japanese Guidelines based on those previously developed for English as a quick way to implement the Guidelines in Japanese teaching. Upon review, however, it seems that the aim of the original ACTFL Proficiency Guidelines was not fully achievable when put into this type of practice. In particular, the "accuracy" category is not only difficult to assess because of the number and variety of its subcomponents, but also because of the difficulty in assessing the relative importance each subcomponent should have in the overall assessment. The difficulties inherent here suggest there might be a better way to organize the problematic aspects of the assessment criteria. In order to do this, it is necessary to return to the components of communicative competence to break down and reformulate these aspects.

Canale and Swain (1980) early on proposed four major components of communicative competence: grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. Grammatical competence and sociolinguistic competence are equally emphasized in their description. Out of this and other early descriptions came one of the more recent frameworks in this area, Bachman's charting of "Language Competence" (1990, p. 87).

1279

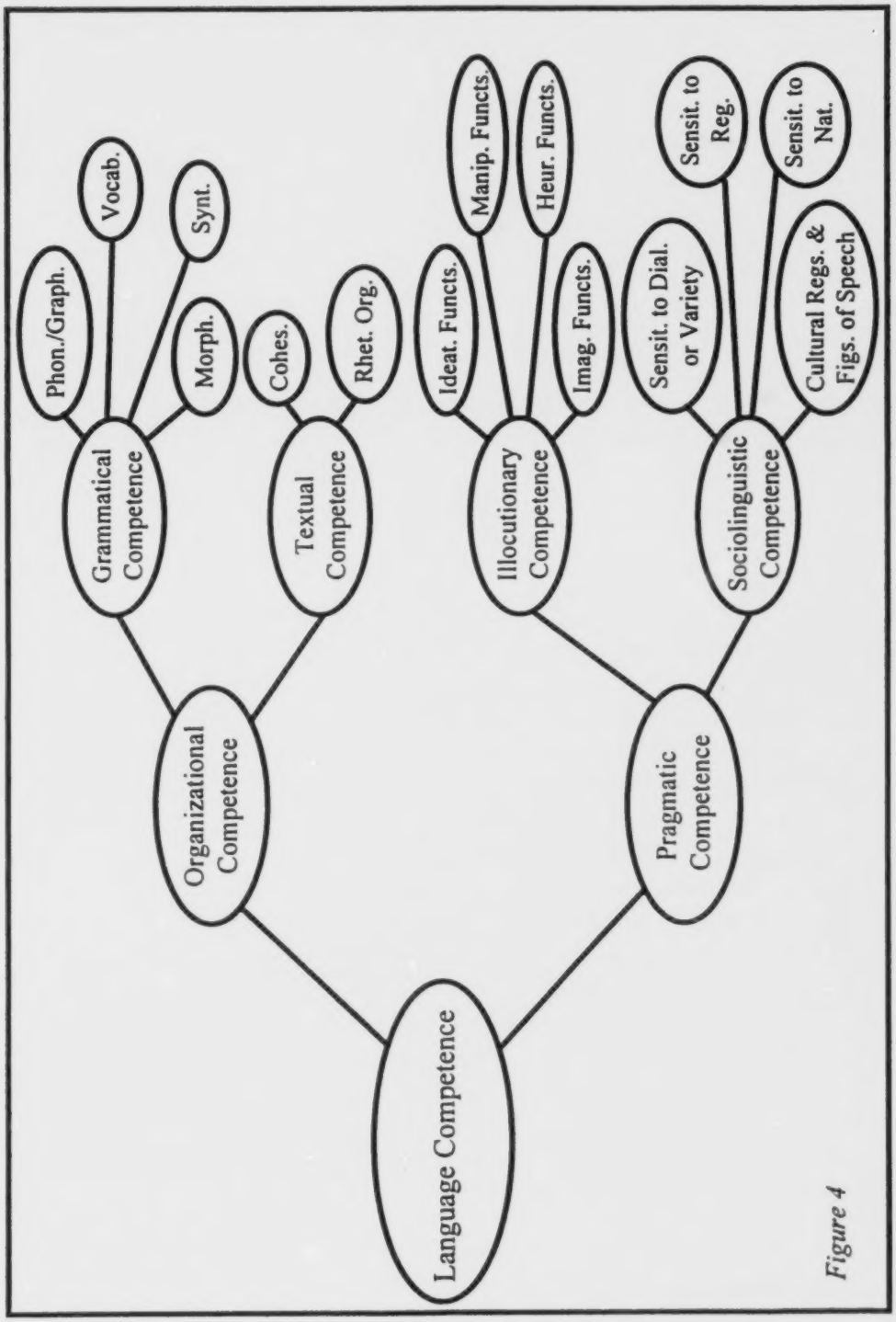


Figure 4

1280

1280

Bachman's components of Language Competence, because the sheer number of categories slows down and complicates the assessment, are a little too subdivided for practical use in university language classes. Nevertheless, it is worth noting that grammatical competence and sociolinguistic competence are different elements of language ability in his proposal. Since the goal of proficiency-oriented instruction is to cultivate communicative competence, assessment criteria should not widely vary from the components of communicative competence. The following chart is a suggested rearrangement of the ACTFL assessment criteria considering the theoretical frameworks suggested by Bachman, Canale, and Swain.

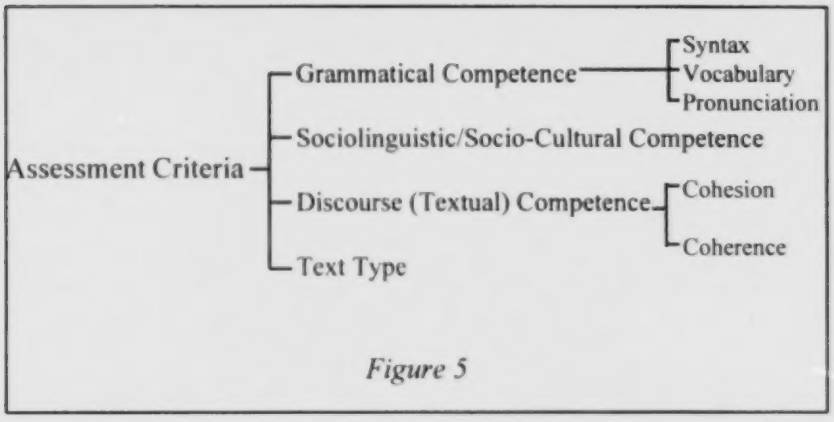


Figure 5

Keeping in mind the purpose of an OPI is to assess speakers' integrative communicative competence, i.e., what speakers can do using the language, an examinee's overall abilities in language use should be assessed through observation of all the subdivided components suggested in the above framework (Figure 5). Three components (pragmatic competence, sociolinguistic competence and fluency) have been separated from the original ACTFL "Accuracy" category (3). Further, in this new division, discourse (textual) competence and sociolinguistic competence are to be assessed independently. This is because pragmatic competence, as shown in Bachman's diagram (Figure 4), spans a number of components of language ability.

Discourse (textual) competence is concerned with whether a given utterance is acceptable in a particular situation. The fluency category has also been removed since the assessment of fluency is actually a decision of whether the learner can use cohesive devices to bind discourse together at a comfortable speed. The use of the word "speed" here, however, does not necessarily mean "fast." Even some native speakers can lack quickness when they are carefully choosing their

1281

words, for example. Therefore, this is obviously a rather difficult criterion to assess. Fluency here is better defined as the ability to communicate with a native speaker without creating an awkward silence, one which would be a barrier to communication. It can be said that fluency is, after all, related to all other factors: selecting vocabulary, forming sentences, listening ability, and experience, as well as personality. It may not be an independently assessable criterion.

One remarkable phenomenon (among many others) which may be regarded as discourse competence is the use of particular vocabulary, phrases, and expressions, or their pointed omission. In Japanese, there is a tendency for "understood" words and phrases to be omitted, and if one tries to create a complete sentence which includes understood words, it sounds rather awkward.

The typical examples of this are the Japanese words for "I." When it is understood that the speaker is talking about him/herself, "I" is unnecessary. Even when the speaker is at a loss for words, "I, I, I, ah ...," is seldom used in Japanese conversation. *Aizuchi* are used instead.

Also, the Japanese word for "you" is one which is almost never used to address a person one is directly speaking with. The name of the person is used. This phenomenon is very closely related to sociolinguistic uses of language and it is not an easy task to draw a clear line between discourse and sociolinguistics in practice.

Japanese society is distinct from many Western societies not in the least because of its so called vertical structure. The correct use of Japanese depends on both whom one talks to and what their status is relative to the speaker. For example, the common Japanese word for "you" is not used to address someone who is superior (stratified cultural ranking) to the speaker, but it is not inappropriate to use it when speaking to someone who is inferior in rank. The pragmatic use of the word "you" is decided by various sociocultural factors. There are also a variety of honorific speech forms which determine relationships among people. The correct use of the honorific forms and informal forms used in intimate relationships is a good barometer for deciding the existence of sociolinguistic competence in the Japanese language examinee. Thus, discourse competence and sociolinguistic competence can in fact overlap in many cases. Nevertheless, the discourse use of language is dominated by sociocultural factors, and this is not reflected in Bachman's diagram (Figure 4).

Accuracy may be the most easily assessable component of all since syntax, pronunciation, and vocabulary use are concrete phenomena. Accuracy in these areas is relatively easy to determine. Grammatical accuracy is also a crucial component in rating communicative compe-

1282

tence. When the level of difficulty is raised and the theme of the conversation becomes more complex and abstract, grammatical accuracy is especially important as a way of communicating without misunderstandings. It is not an easy task to discuss economics and politics, for example, since they require technical vocabularies which may not often be used in daily conversation.

Pronunciation is also crucial. Even if one can produce perfect sentences with cohesion, pronunciation skills may prevent the person from "sounding" like a native, or even being understood. These three components (grammar, vocabulary, and pronunciation) related to accuracy should be grouped under grammatical competence, and each should be given equal weight in the assessment.

In an ACTFL assessment, "text type" is developed at a word level (novice), a sentence level (intermediate), a paragraph level (advanced), and a multi-paragraph level (superior). The test type assessment process may be related to grammatical competence, especially when complex sentences are needed to describe complicated situations. Japanese is an agglutinative language requiring a certain amount of skill in chaining a number of simple sentences together to form a complex sentence. Verb forms have to be changed to accommodate complex structures depending on the content of the sentences which follow them. One may be able to describe what she does during the day, for example, merely by listing simple sentences, but no matter how smoothly one can speak, or how large a vocabulary one has, if the utterances are all short, simple sentences, others' perceptions will ultimately be that such a person's speaking ability is limited. Although, contrary to ACTFL's suggestion (Figure 3), text type should be considered an independent category, grammatical competence may be enmeshed in this category as well. Further research in this area is needed.[5]

The Reliability and Validity of Achievement-Proficiency Tests

Teaching to the ACTFL Proficiency Test in fact creates an Achievement-Proficiency Test. However, such a test is ineffectual if it is found to be unreliable or invalid. Bachman (1990) defines reliability as how much of an individual's test performance is due to measurement error or to factors other than the language ability we want to measure, and validity as how much of an individual's test performance is due to the language abilities we want to measure. He examines reliability and validity in terms of their similarities rather than differences and suggests that reliability is one component which comprises validity. Consequently, if a test is not reliable it cannot provide the basis for

1283

validity.

Many researchers have concerned themselves with exploring the reliability of ACTFL Proficiency Guidelines, specifically the OPI. Clark and Lett (1988) have indicated three major "varying conditions" which will lead to a lack of reliability in the ACTFL OPI (defined here as an achievement-proficiency test when linked to course curriculum):

1. Variation in the types of discourse, topics, and so on broached in the interview producing differences in performance attributable solely or predominantly to the "luck of the content draw";
2. Variation in interviewing technique, including differences in the personality or "style" of the interviewers, that affects performance independently of overall proficiency;
3. For any given interview performance, variation in ratings assigned by the judge(s), either for a given judge relating a particular interview (intrarater reliability), or across two or more judges rating the same performance (interrater reliability).

They go on to suggest uniformity of test content as a way of solving these problems. "The tester could use printed scripts or other aids as a guide, selecting particular types of questions or topics to be broached in the interview"(1988, p. 65). This is most possible for achievement-proficiency tests in a relatively small community. When testers, for instance, discuss in advance the content of the OPI and select a few of the question items or some situation cards that all will use, it helps solve the problem described as a variation in interviewing technique (2 above). As Clark and Lett point out, although an interviewer's style or test personality is not something that can be finitely discussed and prescribed, it can be improved. As to a variation in ratings (3 above), interviewers can exchange recorded tapes of the OPI process to gain interrater reliability, or one interviewer can listen to multiple tapes to gain intrarater reliability.

Although such an improvement requires no small expenditure of time, it is not as impossible as it may appear to be. One advantage is that when the interviewers are also the day-to-day teachers, the examinees will probably not have such a great psychological burden to overcome in taking the tests. They can be relaxed and give certain extended play to their ability. On the other hand, the obvious weak point is that testers may tend not to make an effort to understand what the students' utterances really mean as differentiated from daily classroom observations. This, of course, will lead to a lack of reliability. Thus, interrater rating is a needed component if "same teacher" testing is chosen.

1284

Another difficult problem is validation. As Clark and Lett state:

> proficiency measurement involves determining the examinee's ability to perform in a linguistically and sociolinguistically appropriate manner within a variety of language use situations encountered in real-world contexts external to the instructional setting per se. (1988, p. 54)

The concept of validity is closely aligned with the concept of authenticity. Authenticity questions what real-life behavior is and how people communicate when they are not being tested. When developing an achievement-proficiency test, it is almost impossible to exclusively use real-life situations for testing precisely because real-life is not an overt testing/assessment situation. Therefore, validity becomes a question. Since the goal is to have learners be able to use meaningful language in real-life, once a decision is made to test learners, that automatically limits the range of real-life situations. With achievement in mind, real-life situations introduced in the tests should be related to materials introduced in classroom activities. This does not mean discrete items introduced in class should be on the test verbatim, but rather, those materials should be used as a framework in creating real-life testing items which will derive learners' communicative competence.

The use of achievement-proficiency tests should result in greater reliability and validity than general proficiency tests because communicative competence will be assessed in a relatively restricted range of real-life situations. Test content and testing procedures will focus on a rather smaller area, one based on proficiency-oriented instruction.

Summary

Understanding the symbiotic relationship between the implementation of ACTFL Proficiency Guidelines and Japanese language course curriculum is necessary if, for learners and teachers alike, the value of their merger is to exceed the worth of their constituent parts. In order to reap the greatest possible results from this relationship, it is my assertion that curriculum should be modified so as to embrace a type of proficiency-oriented instruction known as the communicative approach (An approach which aims to cultivate learners' communicative competence).

When the ACTFL Guidelines are combined with course curriculum, the Guidelines lose the characteristics of general proficiency testing since courses will be taught with proficiency testing in mind (mak-

1285

ing them achievement tests). Nevertheless, the aim of proficiency-oriented instruction and the purpose of proficiency testing are still directed at a common goal in that both focus on communicative competence. Consequently, one's achievement will also be one's proficiency. This type of testing is referred to as the amalgam "Achievement Proficiency" test. Because the ACTFL Guidelines mesh well with the communicative approach in the classroom, especially after the ACTFL rating scale is reorganized, the kind of achievement-proficiency test produced should result in greater reliability and validity than general proficiency tests and should serve as a helpful tool to aid in the process of language acquisition.

Notes

[1] This article was inspired by the changes I witnessed in the University of Minnesota graduation requirements. In the early 1990s, there was an attempt to set up proficiency tests based on the ACTFL Proficiency Guidelines and to optimize their implementation for testing proficiency in a number of language programs, including Japanese. Students were required to take a proficiency test in a language at the end of two years of language instruction and to meet the Intermediate-Low level described in ACTFL Proficiency Guidelines in order to graduate. There was a debate in the language departments as to whether or not ACTFL guidelines were useful in implementing such a proficiency test. This article was written as the result of that practical experience to further explore the effectiveness of such implementation.

[2] Over the last three years, after I had already left Minnesota, the members of the Center for Advanced Research on Language Acquisition (CARLA) at the University of Minnesota have developed the Minnesota Language Proficiency Assessments (MLPA) to test students' language proficiency. The MLPA are based on ACTFL Proficiency Guidelines, and the Intermediate-Low level remains the minimal criterion for meeting the graduation requirement (that level would place students in post-secondary intermediate level course sequences in French, German, and Spanish). Asian languages such as Japanese and Chinese are, however, currently not directly involved in this battery (CARLA update: News about second language teaching and learning activities, Summer 1997 Report, e-mailed document, September 1997).

[3] Candline's (1986) definition of communicative competence as "a coming together of organized knowledge structures with a set of procedures for adopting this knowledge to solve new problems of com-

3286

munication that do not have ready-made and tailored solutions" has been adopted for this article.

[4] In this case, the "bottom-up" approach refers to the analysis of grammatical structures while the "top-down" approach refers to more task-oriented approaches.

[5] Another problem in the OPI is that the description of the levels does not take into account the fact that Japanese language levels may not be compatible with those for English. This is supported by the fact that the situation cards used during OPIs cannot be unilaterally embraced by the Japanese OPI. For example, Japanese has many degrees of politeness which have to be put to proper use depending on whom one is speaking with. A perfect command of these speech levels is probably one of the most difficult tasks for learners to attain. Inviting an acquaintance to a party, for example, is placed in the intermediate level of the English role-play cards. However, this is not an equally easy task in Japanese. Consequently, Japanese testers need to create their own situation cards based on Japanese culture.

References

American Council on the Teaching of Foreign Languages. (1986). *ACTFL Proficiency Guidelines*. Hastings-on-Hudson, NY: ACTFL.

Bachman, L. (1990) *Fundamental consideration in language testing*. Oxford: University Press.

Bachman, L, & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal*, 70, 380-90.

Brindley, G. (1989). *Assessing achievement in the learner-centered curriculum*. Sydney, Australia: National Centre for English Language Teaching and Research.

Buck, K. (Ed.). (1989). *The ACTFL oral proficiency interview tester training manual*. New York: The American Council on the Teaching of Foreign Languages.

Campbell, R., & Wales, R.. (1970). The study of language acquisition. In J. Lyons (Ed.), *New horizons in linguistics*. Harmondsworth, England: Penguin.

Canale, M. (1983). On some dimensions of language proficiency. In J.W.Oller (Ed.), *Issues in language testing research*. Rowley, MA: Newbury House, 333-342.

Canale, M, & Swain, M. (1980). Theoretical bases of communicative

285287

approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Candlin, C. (1986). Explaining communicative competence: Limit of testability? In C. Stansfield (Ed.), *Toward communicative competence testing*. Stansfield, Princeton, NJ: Educational Testing Service, 38-57.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

Clark, J., & Lett, J. (1988). A research agenda. *Second language proficiency assessment: Current issues*. Englewood, NJ: Prentice-Hall, 53-82.

Cohen, A. D. (1994). *Assessing language ability in the classroom.* Boston: Newbury House/Heinle and Heinle.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes. (Eds.), *Sociolinguistics*. Harmondsworth, England: Penguin Books, 269-293.

Johnson, K. (1982).*Communicative syllabus design and methodology*. Oxford: Pergamon Press.

Lowe, P., & Stansfield. C. (Eds.). (1988). *Second language proficiency assessment*. Englewood, NJ: Prentice-Hall.

Magnan, S. (1985). From achievement toward proficiency through multi-sequence evaluation. In P. Lowe et al. (Eds.), *Foreign language proficiency in the classroom and beyond*. Lincolnwood, IL: National Textbook, 117-37.

Omaggio, A. (1993). *Teaching language in context*. Boston: Heinle and Heinle.

Rea, P. (1985). Language testing and the communicative curriculum. In Y. Lee (Ed.), *New directions in language testing*. Oxford: Pergamon Press, 15-32.

Richards, J., & Theodore, R. (1986). *Approaches and methods in language teaching: A description and analysis*. London: Cambridge University Press.

Wilkins, D. (1976). *Notional syllabuses*. Oxford: Oxford University Press.

Acknowledgments

I wish to thank Professor Andrew Cohen of the University of Minnesota whose work encouraged me to write this article and who provided constructive feedback on earlier drafts. I also wish to thank Professor Elaine Tarone of the University of Minnesota and other reviewers for their valuable comments.

1288

Author

YUKI JOHNSON, Assistant Professor of Japanese Linguistics and Director of the Japanese Language Program, Department of Asian Languages and Cultures, University of Michigan, 3070 Frieze Bldg., Ann Arbor, MI 48109. Specializations: linguistics, second language curriculum and instruction.

1289

Research Report

Applied Language Learning
1997, Vol. 8, No. 2, pp. 285-293

Voices from the Classroom: Intermediate-Level French Students Favor Extensive Reading over Grammar and Give Their Reasons

Beatrice C. Dupuy
Louisiana State University

By examining the preferences of 49 intermediate-level students of French as a foreign language concerning two classroom activities (grammar instruction and practice, and extensive reading), this study replicates and expands a previous study (McQuillan, 1994) by surveying students studying a different language, and presenting the reasons behind their choice. Similar to McQuillan (1994), students in this study overwhelmingly found extensive reading to be not only more pleasurable but also more beneficial for language acquisition than grammar instruction and practice. Students explained that while reading was fun, interesting, and beneficial for language acquisition, grammar instruction and practice was dull and boring, and its effects small and short-lived.

Abundant evidence exists which suggests that extensive reading is at least as effective, if not superior, to grammar instruction and practice in promoting first and second/foreign language acquisition (see Krashen 1993, 1994 for a review). There is also a great deal of evidence which suggests that extensive reading is perceived as more pleasurable than grammar instruction and practice by first language (L1) and second language (L2) students (see Krashen 1993, 1994 for a review). As extensive reading remains more the exception than the rule in beginning and intermediate foreign language (FL) classes, there has been little opportunity to measure FL students' preferences. To my knowledge, only one study (McQuillan, 1994) has examined how FL[1]

© *1997, Beatrice C. Dupuy*

285290

students perceive both extensive reading and grammar instruction and practice. McQuillan (1994) reported that extensive reading was the favorite among the students he surveyed, however his students did not give any explanation as to why they favored this activity.

By asking 49 FFL (French as a Foreign Language) students to indicate which of these two classroom activities—extensive reading, and grammar instruction and practice—they think is more pleasurable and beneficial for language acquisition, and to explain why, this study replicates, but also expands McQuillan's (1994) in surveying students studying a different language and presenting the reasons behind their preferences.

The Study

The subjects in this study were 49 students enrolled in three experimental sections of an intermediate (4th semester) FFL course at a large public university in the South. All students in this study were assigned to the 15-week FFL class as part of the foreign language requirement set by the university. Although students participating in this study were not randomly selected, none had been informed prior to registration that their section would differ in curriculum, and none requested to transfer from an experimental to a traditional section offered concurrently that semester after the course had begun. All three classes met three hours per week.

In all three classes, students read short stories, poems, tales, short novels, songs, and newspaper and magazine articles that they discussed in small groups during class time. These texts had all been selected by the instructor on the basis of student interests and level of competence in the target language. All students also did self-selected reading outside of class. Less than 10% of the time spent in class was devoted to grammar instruction and practice. However, all these students had been exposed to traditional grammar instruction and practice in other previous French courses taken at the university, and thus had a basis with which to compare extensive reading.

At the end of the semester, after having filled out the traditional evaluation, students were asked to respond anonymously to a short questionnaire, indicating which, between extensive reading and grammar instruction and practice, they found pleasurable and beneficial for language acquisition. In this study, the questionnaire (adapted from McQuillan [1994]) included three close-ended questions (see Appendix) for which students had to choose one of four options. After each question, students were invited to explain the reason of their choice.

PS 1291

The Results

Since a Chi-square analysis of the responses of the three groups found no statistically significant differences between them, they were combined and summarized in the tables below.

TABLE 1
What Students Think is More Beneficial for Language Acquisition

Assigned Reading	Self-selected readings	Combination of assigned readings and self-selected readings	Grammar Instruction & Practice
5 (10%)	4 (8%)	40 (82%)	0 (0%)

Table 1 shows the activity students believe is most beneficial for language acquisition. Reading is unanimously the favorite, with all students (100%) choosing one of the following reading options: assigned readings (10%), self-selected readings (8%), and combination of assigned and selected readings (82%). A Chi-square analysis was conducted to test the null hypothesis that no significant difference existed between the number of students (0) who felt grammar instruction and practice were most beneficial and the number (49) who felt reading assignments (all types of assignments combined) were most beneficial. The analysis confirmed that the null hypothesis could be rejected (Chi-square = 49, df = 1, $p < .001$). These results confirm McQuillan's (1994).

Students favored the combination of assigned and selected readings over the two other reading options. Rather than having only assigned readings or self-selected readings, a majority of the students indicated that they thought that having a combination of both was more beneficial for language acquisition. On the one hand, students explained that while they enjoyed having the opportunity to select their own texts on the basis of their interests and level of competence, the texts selected by their instructor gave them the chance to be exposed to texts they would not necessarily choose on their own, and presented them with a challenge they wanted to take up (*"The assigned readings were definitely more of a challenge, and I liked that. The self-selected read-*

1292

ings were more enjoyable, and provided a more relaxed reading,"; "The assigned readings provided more of a challenge for me and exposed me to genres that I would not generally choose on my own (e.g.: poetry, short stories, etc.). The self-selected readings were sometimes a bit easier, but they made me feel more confident." ; "In this class, I got to choose readings that interest me. Because I chose things that interested me and matched my ability level, I enjoyed reading more and got more out of it. However, it is important to keep assigned readings. With assigned readings, I got exposed to texts that I may not like reading, but that may prove useful later on."). Indeed, assigned readings included a wide range of genres at various difficulty levels.[2] Additionally, students indicated that having the opportunity to discuss and exchange views on assigned readings in small groups proved beneficial too (*"Because the assigned readings were more challenging, and were discussed in groups, they gave me a sense of accomplishment. The self-selected readings were more fun."*). In small group discussions, students had the opportunity to collaborate with more capable peers, and in the process were able to develop their competence in the target language as well as in the way to tackle issues raised in the texts.

TABLE 2
What Students Think is More Pleasurable for Language Acquisition

Assigned readings	Self-selected readings	Combination of assigned readings and self-selected readings	Grammar Instruction & Practice
6 (12%)	27 (55%)	12 (25%)	4 (8%)

Table 2 shows the activity students think is most pleasurable for language acquisition. Here again a majority of the students (92%) favored reading over grammar instruction and practice by choosing one of the three reading options: assigned reading (12%), self-selected reading (55%), and combination of assigned and selected reading (25%). The difference between reading and grammar was statistically significant from chance (Chi-square = 34.306, df = 1, $p < .001$). McQuillan's (1994) results are once more confirmed.

Over half of the students indicated that self-selected readings were

1293

more pleasurable. They explained that having the opportunity to select books on the basis of their interests and level of competence, to read at their own pace, to be able to skip words, paragraphs or pages, or to abandon a book if it was too difficult or did not meet their expectations had been very liberating. (*"Self-selected readings are most pleasurable because you are reading only what you want, and what you are interested in. If something bores you, you have the freedom to move on,"; "The self-selected readings were naturally more enjoyable for me because I was able to choose WHAT I wanted to read. If I found I didn't like it, I could stop reading it, and choose another."*). They also explained that through self-selected readings they had gained confidence in their ability to read in French *("The self-selected readings gave me a great confidence boost. I realized that I could read for pleasure in French."*). In this study, this author had expected a higher percentage than 55% in favor of self-selected readings. To explain a similar result in his study, McQuillan (1994) suggested that "given a choice between reading enjoyable materials supplied by the instructor and having to search for new materials, students [might have been] indicating both their enjoyment of texts supplied and their convenience." (p. 98) The absence of an opportunity to discuss self-selected readings may be another reason. In this study, students indicated that they wished they had had the chance to discuss their self-selected readings with their classmates, as this would have enriched their personal experience with the texts. In both McQuillan's study and this one, students had the opportunity to discuss their assigned readings in small groups, but not their self-selected readings, and this difference alone might explain why students did not overwhelmingly find self-selected readings to be pleasurable. Further study is needed to determine whether this is the case.

TABLE 3
What Students Prefer Overall

Reading	Grammar Instruction & Practice
42 (86%)	7 (14%)

Finally, Table 3 summarizes student preferences, and reading once again comes out ahead with 86% of the students choosing it over gram-

mar instruction. The difference between reading and grammar was statistically significant from chance (Chi-square = 25, df = 1, $p < .001$). McQuillan's results are confirmed one more time.

If, in the future, students were given the choice, a majority of them would chose extensive reading over grammar instruction and practice. Among the 42 who indicated preferring reading, 31% explicitly said in their comments that they found grammar instruction to be dull and boring (*"Grammar is very tedious and boring. Reading is fun."; "Grammar instruction becomes extremely redundant and unpleasant. With reading, there is always something new and exciting to look forward to. It's entertaining!"*). Besides, they expressed a certain disillusion with grammar instruction, for, in spite of all the time spent and efforts made to learn the rules and practice them, the effects have been small, and short-lived (*"Grammar is reviewed and summarily forgotten. If we have a narrative structure to our memory, it makes sense that reading is a more effective way to learn new vocabulary, and grammar."*). Out of these 42 preferring reading, 57% clearly indicated that they found reading to be fun, interesting, and beneficial for language acquisition. Students reported that their grammar and vocabulary improved through extensive reading, and that it happened without them realizing it (*"I've learned more grammar by reading than by direct instruction."; "Reading is more stimulating and less taxing. Grammar is dry, and I find that I learn more from reading because it allows me to see the way French words are used and put together, than by being told and trying to memorize the proper French usage."; "Reading easy literature this semester has greatly improved my ability to speak and read French. It was more beneficial than learning grammar rules or vocabulary words every class period. This makes the class so dull. By reading, I've learned new vocabulary, grammar, and how to read fluently."; "Of all the grammar exercises I have ever done, I have never enjoyed one, but reading is something I like and you learn while doing it without knowing."*).

Conclusion

Students overwhelmingly found reading to be not only more pleasurable but also more beneficial for language acquisition than grammar instruction and practice. These results confirm those previously reported by McQuillan (1994) and expand them in providing the reasons behind student preferences. They suggest that extensive reading of easy and interesting texts, which are then discussed in small groups, is one of the best means that teachers may have to promote a positive

attitude in students and create a more enjoyable environment for language acquisition.

Notes

[1] McQuillan surveyed both intermediate-level ESL and Spanish as a Foreign Language students.

[2] It should be noted, however, that while assigned readings might do some good for college students (with a presumed high level of literacy in their L1 and for the most part, members of the literacy club already), the same might not be true for younger students (especially those who are turned off by books). The dangers of making both children and teenagers (K-12) hate reading through assignments is very strong, as evidenced by research on "reluctant readers" (Bintz, 1993; Duchein & Mealey, 1993).

References

Bintz, W. (1993). Resistant readers in secondary education: Some insights and implications. *Journal of Reading, 36*, 604-615.

Duchein, M., & Mealey, D. (1993) Remembrance of books past ...long past: Glimpses into illiteracy. *Reading Research and Instruction, 33*, 13-28.

Krashen, S. (1993). *The power of reading*. Englewood, CO: Libraries Unlimited.

Krashen, S. (1994). The pleasure hypothesis. In J. Alatis (ed.), *Georgetown University Round Table on Language and Linguistics*. Washington, D.C.: Georgetown University Press.

McQuillan, J. (1994). Reading versus grammar: What student think is pleasurable for language acquisition. *Applied Language Learning, 5*, 95-100.

1296

Appendix

Intermediate Survey
(adapted from McQuillan 1994)

a. Based upon your experience in this class and other foreign language classes you had before, which do you think is most beneficial for your becoming more competent in the foreign language you are studying? CHECK ONLY ONE and explain briefly your choice.

____ assigned readings
____ self-selected readings
____ combination of assigned readings and self-selected readings
____ grammar instruction and practice

Explain:

b. Based upon your experience in this class and other foreign language classes you had before, which do you think is most pleasurable? CHECK ONLY ONE and explain briefly your choice.

____ assigned readings
____ self-selected readings
____ combination of assigned readings and self-selected readings
____ grammar instruction and practice

Explain:

c. Given a choice between reading easy literature and studying grammar, which would you prefer? CHECK ONLY ONE and explain briefly your choice.

____ reading
____ grammar instruction and practice

Explain:

1297

Author

BEATRICE C. DUPUY, Assistant Professor, Department of French & Italian, Louisiana State University, 205 Prescott Hall, Baton Rouge, LA 70803. Specializations: reading as it pertains to language and literacy development, and listening.

1298

News and Views

Applied Language Learning
1997, Vol. 8, No. 2, pp. 295-315

Linguist in Action
Interview with Colonel Daniel D. Devlin

Interviewer: ***Lidia Woytak***

The huge auditorium was briskly filling with young men and women in dark green and blue uniforms who came to listen to the Commandant's Welcome for new students attending the Defense Language Institute Foreign Language Center (DLIFLC). Each time a person walked in, the June afternoon sun made another attempt to get inside, but it was quickly squeezed out by the closing door. These young women and men had not yet developed any preconceived ideas. They were in a zone of questions rather than answers. *What does it mean to be a soldier? What does it mean to be a woman? What does it mean to be a man?* Maybe a quotation from *Wind, Sand, and Stars* by Antoine de Saint Exupéry would answer some of their questions, "To be a man is, precisely, to be responsible."

In today's world of subliminal profit-oriented messages, more than ever these young people seemed to be looking for someone who would set standards for them and who could at the same time serve as an example. They needed a person who would be capable of walking in their new shoes. They needed an individual who would care so much about them and would not be afraid to sound or look unpopular or untrendy. Over the years, students—assembled within the walls of this old Tin Barn—have listened to different types of commandants: Some of them liked to emphasize the academics, some the military routine, while others would rather focus on infrastructure or physical fitness. Now amidst rows of female and male soldiers, sailors, airmen, and marines walked a commandant to whom leadership, personal responsibility, and concern for his people were daily guideposts.

Talking to several hundred students about all aspects of their life, Colonel Daniel D. Devlin told them that he cares about each one of them personally and professionally. While instilling a sense of responsibility in his soldiers, Colonel Devlin did not hesitate to say, "I care about you." Further he informed them that if they have any questions or concerns, they should follow the chain of command but if all else

1299

fails, they can always talk either to him or to the Command Sergeant Major. He continued, "Because I care about you, I want you to be responsible." While warning about drinking or substance abuse, Colonel Devlin tied the concept of responsibility to an awareness of choice. He stated, "Know what you're doing and know the consequences." In his presentation he heightened the sense of responsibility in the new students by asking them to look at it as a refusal to make bogus excuses. Further, in the spirit of St. Augustine's *Confessions,* "all because we are ashamed to hold back when others say, 'Come on! Let's do it!'", Colonel Devlin spoke against a weakened sense of responsibility under peer pressure.

Colonel Devlin's holistic approach has placed the Institute's students at the center of his operations. Apart from instilling behavioral standards, he is striving to provide an optimal learning environment for each DLIFLC student. In accordance with the United States Army's vision of the soldier of the twenty-first century, the Commandant is trying to equip every student with a computer so they can develop and enhance language skills in and out of the classroom. Colonel Devlin believes that computers can be used to help transform students into quality linguists who can solve language problems by themselves anytime and anyplace. His holistic approach to student life-style at the Institute works. General Dennis J. Reimer, Chief of Staff, US Army, wrote recently, "The Defense Language Institute performs an invaluable service by providing our Nation with superb linguists." (1 Oct 97)

Since his arrival in February 1996, Colonel Devlin has been working within and beyond the Institute to prepare military linguists for the new Millennium. His actions fully reflect military doctrine as outlined in two seminal articles titled "Joint Vision 2010" by US Army General John S. Shalikashvili (*Defense*, 4/1996) and "The Quadrennial Defense Review" by Secretary of Defense William S. Cohen (*Defense*, 4/1997). The authors of these articles call on leaders of US educational and training programs to "prepare joint warriors to meet the challenges of the future battle-space" through employment of new technologies.

Many of his projects involve players beyond the Institute. Right now in cooperation with commanders in the field, he is reviewing the training sequence of DLIFLC graduates. At the Command Language Program Manager (CLPM) Seminar in May 1997, he asked the linguists:

* Are we serving your needs?
* What can we do better?
* How can we improve our program?

1300

In these times of political correctness, Colonel Devlin's willingness to praise and to assess various projects at the Institute is refreshing. Following his presentation at the 1997 Annual Program Review during which he pointed out what needs to be done in the areas of curriculum and technology, a representative from one major agency commented, "I truly enjoyed listening to Colonel Devlin's speech, he cares so much about the Institute and the students."

Graduating from North Dakota State University in 1969 as a Distinguished Military Graduate with a Bachelor of Science in History, Colonel Devlin subsequently obtained a Master of Arts in National Security Affairs specializing in Soviet/East European Areas from the Navy Postgraduate School and a Master of Arts in International Relations from the University of Southern California.

Colonel Devlin is a graduate of the Armor Officer Basic Course, the US Marine Corps Amphibious Warfare School, the Defense Language Institute Russian Language Course, the US Army Russian Institute, the US Army Command General Staff College, and the US Army War College. He is also a graduate of the US Army Foreign Area Officer (FAO) and Psychological Operations Officer Courses.

Colonel Devlin was commissioned an Armor Second Lieutenant in 1969. From 1969-73, he served in a variety of armor command and staff assignments, and as the Aide-de-Camp for the Commanding General in the 1st Infantry Division (Forward) in Germany. From 1973-76, he served in armor command and staff assignments in the 1st Cavalry Division at Fort Hood, Texas.

In 1977, Colonel Devlin was selected for Soviet/East European FAO training. Following civilian schooling, attendance at the US Army Russian Institute, and attendance at Command and General Staff College, he served from 1983-88 in the 66th Military Intelligence Group/ Brigade in a variety of FAO assignments. Colonel Devlin commanded the 6th Psychological Operations Battalion (Airborne) at Fort Bragg, North Carolina from 1988-90.

In Operations Desert Shield/Storm, he served first as the Deputy Commander of the 4th Psychological Operations Group in Saudi Arabia; then as Liaison Officer to the American Embassy, Cairo; and finally as the Liaison Officer to the Joint Special Operations Command, supporting operations in the Persian Gulf region.

In June 1992, following attendance at the US Army War College, Colonel Devlin was assigned to the Pentagon as the Chief of Psychological Operations and Civil Affairs for the Joint Staff, where he served until February 1996 when he assumed command of the DLIFLC and Presidio of Monterey.

1301

What is your vision of DLIFLC?

The most important vision for the future is to ensure that we not only have continuing strong language education but that we fully integrate technology into our language instruction. Specifically we can apply new technologies in computer-assisted and adaptive training to ensure the students have maintenance, sustainment, and enhancement capabilities that computer training has to offer.

At the same time, I want everyone to understand I fully believe that the best language learning environment is created by a language teacher in a classroom with a group of students. Technology integration is needed to support language education, not in any way is it a replacement of a language teacher.

What are your main actions to achieve this vision?

I would hope in the future to be able to get additional computer support. I'd like to see every student working with a computer, but that requires additional budget funds.

What has DLIFLC accomplished in achieving this vision since you came on board in February 1996 and what are the major gains?

The technology staff in the Curriculum and Faculty Development Division is doing a better job of providing the computer-adaptive instruction. We've done some real work in Serbian-Croatian and Russian, and currently, we're preparing some additional instruction in Arabic. It's difficult to produce computer-adaptive instruction in all languages immediately because we have insufficient staff and funding.

So how are we going to fill the gap between what has been done and what still needs to be done?

We're going to keep working to get additional funding wherever we can find it. We will apply to TRADOC for funding in special areas. We'll ask staff involved in distance learning to provide additional training or perhaps to conduct test projects to see whether computer-

1302

enhanced learning is a better way of providing instructional support to students.

What is your recipe for an optimal institutional change?

The optimal change would be to have computers issued to all students upon their arrival here so they would begin language instruction with an understanding that computers can be used to remediate, maintain, and enhance language capability. Thus, by the time they leave the Institute, they would be already well adjusted to using the computer for language development.

The Joint Vision 2010 calls on the Armed Forces to do a better job in closing the looming gap between the military and the society at large. What measures are you taking in getting the Institute closer to the American society at large?

That's a pretty good question. Right now recruiting commands across the United States are doing a better job of getting in touch with high school and college instructors so we can host their visits to the Institute. They come and look at the kind of instruction that we provide as follow-on instruction for students learning foreign languages today either at a high school or a college level. Additionally, every spring during our Language Day open house, we invite students from California and the neighboring states to come visit us and see what we do here on a firsthand basis.

What is the Institute's committment to the Joint Vision 2010?

In the spirit of the Joint Vision 2010, we enhance our relationship with the civilians by reaching out to teachers, students, and recruiters. We provide information about language training at the Defense Language Institute Foreign Language Center. Currently, we are updating the video about the Institute.

In the quest for operational excellence, the Joint Vision 2010 calls for reliance on new technologies. How do you envision the potential of new technologies within the DLIFLC curricula?

1303

First, computer technologies can provide support to students who are trying to remediate, maintain, and enhance their language abilities. If we could tie doing daily homework on a computer to students' understanding of how to improve their language, then the computer would become a support system for them wherever they go.

Second, we are already using technology for distance learning. We lead the Army in the use of distance learning for language enhancement and maintenance to our alumnus via video teletraining and video teleconferencing. Third, we now have desktop video capability that allows us to reach out to other language students.

What role do the academic journals, Applied Language Learning *and* Dialog on Language Instruction, *play in bringing about an institutional change?*

Any institution of the DLIFLC caliber should be able to provide professional materials pertaining to the ongoing tasks and challenges. Our two professional journals are a cut above the kind of information we provide in other ways. They have different purposes. *Dialog on Language Instruction* is tied more to the teacher here at DLIFLC and offers information on the kinds of instructional techniques and technology advances that can help in our classrooms. *Applied Language Learning* is more appropriate to foreign language teachers across the country and around the world. We exchange ideas with researchers and foreign language teachers on inventive ways of improving language instruction.

Do you think that such an exchange should include military and government language schools as well?

Absolutely. These schools offer varied kinds of instruction from what we do. Nevertheless, it's important that we reach out to civilian academic institutions as well because some of them may have tremendous advances in teaching foreign languages, particularly using computerized instruction.

How could the role of the academic journals expand?

1304

We're expanding it by ensuring the journals are offered to the kinds of institutions that need them and those language professors, as well as testing and evaluation researchers, have an opportunity to publish in our journals. We are also offering the Institute as a place where researchers who are conducting foreign language learning studies can come and use us as a test site.

In your monthly "Commandant's Welcome" to new students, you said you talk straight about the consequences of drugs, sex, and alcohol abuse. Why do young people need such straight talk?

I am concerned that young adults in our society, mainly between the ages of 17 and 25, the bulk of our population here, don't get a proper warning. Too often leaders couch their words when they talk about these subjects to them. These young service members have signed up to wear the uniforms of their country regardless of which service they chose and have signed up to be willing to defend this country. Thus, we owe them straight talk so that they understand exactly what we're talking about. My interest in discussing these issues with them is that they learn to be responsible and that they learn to accept the consequences of the decisions that they make.

You say the student is priority number one. Could you elaborate what that means to you?

That's a little misquote of what I say. What I really say is that the student is our first customer. As a priority, the students and the faculty together are priority number one because without them we don't have an institute. However, the customer, our first customer, is the student because without him or her we would not have a reason to exist. In the past I'm afraid that we didn't look at students as customers: We only saw them as a product that we provide to our ultimate customer, the commander in the field who uses our students. We need to focus on the students, making sure we're providing for their needs from an instructional base as well as from the operational support base. If we don't look at them as customers, then we're missing the value of what they offer us.

You tell the incoming students you have an unwritten contract with them. Could you reveal its contents?

2.1305

Certainly. The unwritten contract simply says that my part of the contract is to offer the best civilian instructional team that I can for teaching them foreign languages and their part is to do their best in learning foreign languages. That's what I tell them, I am not asking them to be superhuman: I am asking them, "Just do your best and I'll do my best to provide the best teachers possible."

Right after your "Welcome," I talked to two new students who said how greatly they appreciated your presentation. Why do you care so much about these young people who are our students?

These young people who come to the Institute are the future of our language capability as a Department of Defense and as a government. Many of them will use this language as long as they are in the military and will use it whether they go into other government service or whether they go into a commercial field that requires a foreign language. We don't do well enough in this country in teaching foreign languages at the high school and college levels and since DLIFLC provides about 13 percent of the foreign language instruction above high school level, we've got to do it right. These young people are important. Moreover, my concern for these young service members just comes down to the fact that I am a leader in the military and leaders are primarily responsible for ensuring that soldiers, airmen, sailors, and marines are taken care of. That's my interest.

President Harry Truman kept a handmade wooden sign that read, "The buck stops here." on his desk at the White House. In a similar vein, you alert our students to wooden signs bearing your name and the name of the Command Sergeant Major at the entrances to the Installation. You tell them that the signs are there to remind them you are responsible for their welfare at the DLIFLC. Why are you so direct about your responsibility?

Because they need to know that if there is a problem and I don't know that the problem exists, I can't fix it. I want them to understand that if they don't get their questions answered any place else, they can ask me. There is a chain of command. Thus they have the opportunity to go through the chain of command which consists of noncommissioned and commissioned officers, and me. Nevertheless, I constantly remind them that if all else fails, they can refer to the two people whose names are on the signs at the entrance: One is mine and the other is the Command Sergeant Major's. Both of us want to ensure we know about

9306

problems so that we can get them fixed.

Many leaders of educational institutions view their students or service members as adults and consequently do not warn them about perils facing them in and out of the classroom. Wouldn't it be easier for you to take a similar stance?

I don't think leadership is about taking the easier route. It's about ensuring that the people whom you're trying to lead understand their responsibilities and understand that when you take responsibility, you make decisions, and then you live with the consequences of those decisions. That's what I'm trying to tell them. We have a responsibility to each other, we have a responsibility to ourselves to ensure that we have thought about the consequences of our actions before we take those actions.

What are DLIFLC efforts in the areas of curriculum and faculty development? Are we doing enough in those areas?

Although we're doing a better job in teacher training now than in the past, we don't have enough follow-up training. In addition to our initial training for newly hired teachers, we need to offer them viable follow-up training throughout their careers at the Institute. On the side of curriculum development, many of our curriculums are overdue for revision—the average age is over 10 years with some over 30 years old! Unfortunately, revision is a funding issue. We are providing curriculum development reviews followed by curriculum reorganization wherever we can afford it, on a priority basis.

On many occasions you expressed concern over high attrition reaching its peak towards the end of the first trimester at DLIFLC. How can we reduce attrition rates while increasing quality of our students?

If students were doing homework on computers, they could be tested on a daily basis. Such testing would indicate instantly when they start getting behind and it would identify the areas of their weaknesses. Attrition is all about students getting behind and being unable to finish with their classmates. If we can promptly identify where the weaknesses are, then we can perhaps understand our attrition rates better and hopefully decrease them. I am greatly interested in computer-adaptive training because it would identify weaknesses, provide diagnosis, and relevant instruction.

Occasionally students who have language learning difficulties end up talking to you. What do you learn from these conversations?

Students who end up here have been recruited into their services and the recruiters are often unable to give them an accurate picture of how hard they're going to have to work. So you end up with three types of problem students.

The first one is a student who is trying very hard and is motivated and is working at his/her maximum capability, but has been placed in a language that is too difficult for him/her. In this case, we try to find a match where the student gets assigned to a language that he/she is capable of learning.

The second one is a student who signed up for a foreign language and didn't understand how difficult it would be and is not motivated to put in the effort, in other words, to try his/her best. In this case, we try to come to an agreement with the student that he/she might be better off and better suited doing something else in the military.

The third one is a student who simply does not have the capability for learning a foreign language. It is not cost effective or efficient to spend a large amount of time trying to teach a foreign language to someone whose capability for learning it is low. Again we try to explain to such students that learning a foreign language is just too hard for them and they might be better off learning a different specialty.

I suppose there is a fourth category. Such a student is someone who for either mental or physical reasons simply can't stay in the military and consequently has to be dropped from a foreign language program.

What has made you conclude that no one is a failure except those who do not try?

In the past there was a tendency to send students away from DLIFLC who couldn't make it in a language with a message that they were failures. Although foreign language learning can be too difficult for some people who have been sent here, I don't want them to think that they are failures because they haven't been able to learn it. I want them to come to an understanding that their time and effort would be better spent doing something else in which they would be happier and which would be a better match for their capabilities. The unsuccessful student, who tried his hardest, is not a lesser person than a successful one sitting next to him. He just has less capacity for learning the language: He has skills in other areas that he should take advantage of.

AU1308

You encourage the new students to focus on the essentials of language acquisition. You say, "The secret is the key." Can you elaborate on your statement?

You simplified the quotation a little bit. Using a variation of an English poem written in substandard English and difficult to decipher, I tell the students that once they understand how it was written, then they can read it. So if they understand the keys to a foreign language, for example, once they learn the Cyrillic alphabet for Russian, they can understand how to read it. But without understanding the alphabet, they can't read it. The same is true of an Arabic script or characters in an Asian language. Without the ability to decipher script or characters there is no way for students to register the written representation of that language. Once they learn the keys, then they can study the language in its written form.

Lieutenant Colonel Roderic A. Gale, DLIFLC Associate Provost and Dean of Students, in his article "Computer Roles in Language Teaching and Learning: Let the Dialog Expand Robustly!" published in the June '97 issue of Dialog on Language Instruction ***describes your concepts of computerized training and computerized homework. Could you comment and elaborate on these two concepts?***

Through computerized training, students can remediate their language ability, sustain, and hopefully enhance their skills. The difference between computer-assisted training and computer-adaptive training is that with computer-assisted training you basically have to work yourself through the levels to get to the level you are capable of learning. In computer-adaptive training, the computer helps you find that level, so you don't waste time on skills that were too simple for your capability. Therefore, we need to focus on computer-adaptive training.

The computerized homework would allow us both to have a student more easily do the homework in a fashion that would allow us to centralize its collection and, therefore, let the computer analyze a group of students and their progress. Imagine students doing their homework each night and a language teacher asking the computer each morning, "How did my students do?" The computer would identify a certain concept a few students didn't grasp or the segment of instruction no one in the classroom understood. The teacher then would identify the portion that must be readdressed and solidified in students' minds before they can move on. We currently don't have a method of tracking homework of each student individually and as a group. Computers

309

could help us in this regard.

Do you believe students could use computers to practice simple tasks without the assistance of an instructor? How could that be done using computer-based training?

Students could use computers to practice simple tasks such as pronunciation, spelling drills, or even memorization of vocabulary. They could practice language on computers on their own time rather than wasting time in the classroom where only one student can perform at a time. Thus, the students would be doing individual tasks on their own time and would use computers as guides. On the other hand, collective tasks would be done in the classroom where students could discuss various topics and the teacher could present new learning techniques and new concepts rather than spending the classroom time on repetitive tasks.

In the 80's, a significant number of graduates ended up assigned to non-linguist jobs. In reaction to this situation, one of the DLIFLC graduates wrote in the 1987 issue of Dialog on Language Instruction, *"It does not make sense to train an individual for a year then send him off to wash tanks. That is a waste of resources." (4/2, p.66) Has the situation improved in this area?*

Although the assignment of linguists after graduation from the DLIFLC doesn't fall within my responsibility, I address this issue wherever possible. If Army, Navy, Air Force, and Marine Corps personnel commands could early identify requirements for linguists, then the assignments for them would be easier.

The case where someone learns a foreign language and then ends up washing tanks sounds like an exception. There are language graduates who believe that the only thing they'll have to do is use their language. That simply does not fall in line with being a soldier, sailor, airman, or marine, because there are other tasks that always are required. But it is important that linguists employ language skills in a useful way and are also allowed both on duty and off duty time to practice, maintain, sustain, and enhance them. I emphasize this part in discussions with commanders in the field when I go and visit them.

Under your leadership, the annual Command Language Program Manager (CLPM) Seminar has grown from a few dozen to over 200 participants. At the May '97 Seminar presenters and roundtable discussants made numerous specific recommendations pertaining to

1310

operations of language programs. Could you bring us up to date on the follow-up actions?

I don't think the CLPM Seminar has grown because of my leadership, but rather it happened during the time I've been here. Certainly I believe in the program and I encourage people to participate in that seminar. We follow up on participants' recommendations and particularly the roundtable discussions by collecting all comments and providing feedback to them in the field. This year we have done a better job than ever in following up on the comments and suggestions. We do a follow-up dialog after every seminar or conference by the use of computerized communications, e-mail specifically. We just concluded a conference on technology integration and I promised the participants to provide feedback on every comment.

The concept of linguist lifecycle management was extensively discussed at the CLPM Seminar. Could you describe this concept?

Linguist lifecycle management ensures that a linguist who has been taught a foreign language has the opportunity to maintain his abilities and hopefully enhance them over time. For example, we create an opportunity for a language graduate of a basic program to take an intermediate course or an advanced course in the future. Additionally, our graduates have the opportunity to get further language instruction via the video teletraining system which is tied to the Institute.

What is DLIFLC's role in the linguist lifecycle?

DLIFLC's role certainly is to provide continued distance instruction to our graduates and to reach out wherever possible for remediating, maintaining, and enhancing language abilities. Currently we are developing compact discs and desktop videos for language enhancement.

In the opening speech of the CLPM Seminar, you stated you would like to see assignment of the DLIFLC graduates directly to the linguist jobs so they would not lose language skills in the interim. Do you see it happening?

I'm offering my opinion in an ongoing discussion on sequencing of language training. It seems logical to schedule a common block of advanced individual training for linguists in the intelligence field before they come to DLIFLC. Once they have completed training in the

U.S. 1311

intelligence field, they could get their language training. Afterwards they could go either directly to their unit for on-the-job training or they could go back to professionalized training at a centralized location where they would learn to use linguist-related equipment. Today the problem is that for between 18 and 26 weeks of training there is little or no language training taking place during the advanced training. Therefore, they tend to lose some of their language skills.

At the Seminar, some linguists complained about too much reliance on multiple-choice assessment of language skills. Further they asked for development of more diagnostic tests and maintenance materials. To what extent has DLIFLC addressed these issues?

We have an important mission, of course, of providing evaluation of standardized testing across the field for linguists. In centralized testing programs, a multiple-choice test becomes an easier method to determine language capability, to grade, and to get results back to the linguists. We have a good reliability record of determining linguist capability from the multiple-choice test results. We do have the ability with our new desktop video to conduct face-to-face language evaluation in an interview process with a native linguist. Since an interview is a subjective test, we have to tape them in case of any difference of opinion over the rating. If needed, a second or a third native linguist can listen to it and give us an additional rating. Thus, an interview may take a lot of time.

What should the DLIFLC be doing to bring the Command Language Program Community closer together?

We have the network on the Internet called LingNet that provides the ability for linguists and Command Language Program Managers around the world to exchange ideas and information. We currently are adding information to LingNet every day and we are receiving more hits on that network every day. Last time I checked we had approximately 2,000 inquiries per day on LingNet. As long as we continue to update our information so that we're exchanging information among Command Language Programs, we are supporting the field very well. For example, the language program at the Medina RSOC offered the forms to keep track of linguists training on a system that we put on LingNet. These forms can be used by any language program manager around the world. We also provide support to the field by encouraging language managers to share with us their ideas and by ensuring that we

312

have looked at them to determine their relevance.

During your opening speech to the CLPM Seminar you stated you wholeheartedly support Worldwide Language Olympics. Why do you consider Worldwide Language Olympics to be important?

Because it's a fun way for linguists to practice language capability. Worldwide Language Olympics has grown, as you said the Command Language Program Manager Seminar has grown, to the point where there are now regional competitions to determine who gets to come to the competition that is held here for the worldwide competition. Such an enthusiastic response tells me that linguists are spending more time practicing their skills in order to take part. Any means that we can come up with to get linguists to practice their job is progress, in my mind, towards making them better linguists.

What else is DLIFLC doing to enhance careers of linguists?

From a career management standpoint, we participate in a discussion of how better to assign linguists, to provide additional training to them in the future, to get them back to intermediate and advanced courses. We participate in those ongoing discussions, but we don't have the responsibility for future assignments. We do offer our advice and guidance for linguists and we certainly encourage them to come back and get intermediate and advanced training.

How would you describe linguists' readiness for duty?

It depends on where they go. I've visited several sites recently and as long as the commanders in the field are requiring language maintenance on a regular basis and in most cases requiring it as part of the duty day, I think that is taking care of linguists and assisting their readiness for duty.

How does a linguist career compare to other careers in the military?

The linguist skill is perhaps best compared with a technical skill in aviation. Like in aviation, you have to go to school for quite a while to learn it. You have to continue to work at it in order to maintain it and you can be rewarded for advancing your skill level. Since the language

AE1313

skill is one of the most difficult to achieve and to maintain, linguists have to be required to continue to work at enhancing their skills just as aviators are. I see no reason why linguists shouldn't be treated with the same amount of care. From an emotional and mental state, there's very little training in the military that's more stressful than learning a foreign language at the pace we teach it.

Do you see the challenges facing DLIFLC as opportunities or problems?

Opportunities are almost unlimited. Look at the commercial application towards language learning today: Almost every corporate leader in this country needs people who can speak a foreign language to send around the world to increase their markets. Now we're talking about commercial applications and that the people we're training here are military and government linguists. We can solve our problems both in the military and commercial world by being able to understand, comprehend, and communicate with others in their language, not just in ours. The foreign language programs at DLIFLC provide linguists for the government and subsequently for the commercial sector.

Over the years, we all went through different management waves such as management by objectives, search for excellence, one-minute manager, thriving on chaos, etc. Today, management guides tend to emphasize attention to detail and enthusiastic approach to work. Is there a particular school of management that you subscribe to?

Good leadership and taking care of your people. Every leadership concept or every management concept that I have seen offers something different and therefore has to be looked at on its own, not directly compared with other systems. The problem with almost every management concept is that it usually is applied in a very strict way without any flexibility of application.

The most recent wave is total quality management. Once again people come up with absolute checklists and frameworks that are supposed to apply to every situation. Since the concept of total quality management fits every organization differently, it must be applied flexibly. Regardless of management style, good leadership and taking care

1314

of your people will always result in successful mission accomplishment.

There are many language teaching and learning teams at the Institute. What ingredients make a team work?

The ability to work with each other and a goal of attaining an end result—which is your students learning a foreign language. If you have good leadership within the school, with the chairpersons, and within the team, and all the people work together—then you're working towards a common goal.

In your career you have encountered many good and bad managers. What are the common qualities of good managers?

Understanding that there is a goal to be accomplished, providing leadership and guidance, providing rewards. In my mind, positive motivation techniques are much better than negative ones. And taking care of your people. When you take care of the people who work for you—they, in the end, are interested in working hard, are interested in success.

If an inexperienced or a "would-be" manager asked you, "How can I become a good manager?" what would you say?

I would say, "Become a leader." There is a constant dialog that goes on regarding the difference between management and leadership. They are probably very similar in the military. Leadership is ensuring that all are working towards achieving your goals. Management is simply using the assets that are available to you in a way that allows you to accomplish that goal. But I would say to a manager, "If you don't understand that you've got to take care of your people and manage your assets in a way that allows you to make the most out of everything you have, then you probably won't be a successful manager."

How do you compare student life at DLIFLC today to that of almost two decades ago when you were a student here?

1315

First, we work the students harder today, but, at the same time, we're providing better instruction. We teach the language teachers more about their jobs and the results show that they are more professional. We review and modernize the curriculum in almost every language not because we're doing poorly in instruction, but to make sure that we fit today's growing technological orientation and that we have current cultural and linguistic materials. You're not going to have state-of-the-art language training unless you update your curriculum.

How do you go about ensuring that the Institute makes progress? We ensure that we have trained teachers and we demand the most of our students. Second, we have established a routine here today that ties together the physical fitness capabilities of the student and the mental challenges of the student. The routine is set up in a way that enhances language training. During the course of the day we apply those types of training techniques that are best for that time of day. We start out by teaching new concepts in the morning, which is the best time to learn. In the early afternoon, to keep students alert, there are discussions and debates during which they can be challenged by what's going on. In the late afternoon, they get physical fitness training which supplies oxygen to their blood, blood into the brain, and charges them back up for the homework at night. We're doing a much better job of looking at the whole student rather than just trying to focus on language at the exclusion of everything else.

How do you imagine the student life-style at DLIFLC five, ten, fifteen years from now?

I'd like to see that we'd continue to offer professional training to our teachers so that they would be empowered to offer better and more challenging language training to our students. Second, I'd like to see that we offer our graduates better means of language maintenance and the sustainment capabilities through Internet access. Thus, wherever they would go, they would literally be tied back to us.

What study habits have worked for you?

Hard work. Long hours. The most important study habit for me or any student is the one that is most comfortable for him. You'll find students who do very well with music playing in the background and

1316

there are others that have to have absolute quiet. It's very important that students have the awareness that they can create a study environment that suites them the best. To allow reinforcement, we do the best to ensure that students who are in the same language room together, they have a quiet place to study and that they can have music or television whenever they take a break.

Additionally they have communications available to them, which is important towards mental well-being. They can talk to their friends and loved ones back home. That wasn't always the case. Today students can have telephones in their rooms and can talk to friends and relatives and loved ones back home.

What was the major difficulty for you in learning Russian?

The major difficulty for me was memorization and retaining Russian vocabulary, then trying to keep it separate from languages I had already studied. It would have been easier if I had learned it when I was younger. I wish that we could teach students foreign languages in kindergarten; they would probably learn it without accents.

What were the major barriers for other DLIFLC students? What fears did they experience then?

The hardest part for any American student is that in order to understand the grammar in a foreign language they have to understand the concept of cases, which we don't cover well in our English curriculums. As a former high school teacher, I had to teach English grammar to younger students so that they could then understand the concepts of a foreign language grammar. Perhaps grammatical misunderstanding is the hardest part for most students when they come here. Certainly new grammatical concepts were difficult for students and they expressed fear at ever being able to learn them. Matching vocabulary with grammatical skills and then putting it into practice was always an awesome challenge.

Tell us about your best teacher at DLIFLC.

The best teacher I had was a teacher who was capable of explaining

difficult concepts both in the foreign language and in English. A teacher who understood the English grammar rules as well as the foreign language grammar rules. It happened to be a young teacher who had learned Russian as a foreign language and knew a couple of other languages as well.

Describe to us your best day in the Russian program of DLIFLC.

Graduation! Seriously, it was any day whenever learning the language was combined with having a good time with classmates and the instructor. Nothing should be a drain on your emotions and your physical state. If you enjoy what you're doing, you create a better place to learn.

Following your graduation from the Russian language program, what did you do to enhance your language skills?

I went directly to a job that required the language and got further language instruction.

In what ways did Russian, a foreign language, help you to understand better your native language and culture?

It's more a case of recognizing the fact that when you speak Russian to a Russian native you feel like a visitor to America trying to speak English to us. Most people want to help you communicate in their language. Graduating and speaking Russian and having people compliment me on my Russian capabilities was a pleasant surprise after having been corrected so many times in the classroom.

How have language skills helped you to become a better Foreign Area Officer?

It's very difficult to be a successful Foreign Area Officer without understanding the language, culture, history, social backgrounds, and religions of the area. If you can't visit the area and read the things that are around you and speak to the people in their language, then you probably can't be very successful.

1318

How does it benefit the Institute to have a former student command it?

Understanding what the classroom is all about. If you have not sat in a classroom seven hours a day studying a foreign language and then done three to five hours of homework a night five days in a row for a year, then it's difficult to understand the stress of being a language student at DLIFLC. I have sat through that routine. A student can't come to me and complain about situations in the classroom without me understanding what he/she is talking about. That's probably the most important advantage of being a commandant after being a student.

What is the question I forgot to ask you?

If I like the job I'm doing now. I love the job I'm doing now because I'm dealing with language students who are here to learn in a system that I've already gone through. I really wish that I could spend all my time with the students because it's exciting and rewarding. I also have other duties and some of those are not always as much fun as working with language students.

1319

Reviews

Penguin Readers. Series Editor: Derek Strange. London: Penguin Books. (http://www.penguin.co.uk).

Reviewed by PAUL NATION
Victoria University of Wellington
New Zealand

The Scheme

This series of graded readers seems to have appeared in March 1995 and since that time the number of titles in the series has rapidly grown to well over a hundred. The titles include adaptations of well-known best-sellers like *The Boys from Brazil, Free Willy, The Client, The Pelican Brief, Babe* which have appeared as very popular films, classics like *Dracula, Black Beauty, The Picture of Dorian Gray* and *Treasure Island*, and humorous texts like *The Book of Heroic Failures, My Family and Other Animals* and *The Darling Buds of May*. All titles are fiction and are largely adaptations of existing texts rather than original stories written for learners of English.

The striking features of the series are the well-known titles (see the film, read the book), and the extremely attractive covers (a scene from the film wherever possible) with the familiar orange penguin rib and logo. They look like ordinary unsimplified novels, and it is not difficult to imagine a teenage learner of English as a foreign language looking "cool" reading an English novel while travelling on the subway.

The books are controlled for vocabulary and grammatical structure. Their length is also controlled with the texts becoming longer at the higher levels. There are six levels in the series, ranging from a vocabulary of 300 words at Level 1 to a vocabulary of 3,000 words at Level 6.

The grammatical structure grading scheme is outlined in promotional publications and for each level is briefly described in a note in each reader. The vocabulary grading scheme is not described and does not seem to be readily available. According to promotional material "vocabulary is kept strictly within the headword count at each level." This is reassuring, as there has been a tendency for some series

1320

to abandon strict vocabulary control.

The range of titles includes those attractive to both younger and older learners, although it seems that readers at Levels 1 and 2 are aimed mainly at younger learners.

TABLE 1
Total Vocabulary Increase at Each of the Six Levels in the Penguin Readers Series.

Level	Total vocabulary	New words added at each level
1	300	-
2	500	200
3	1,050	550
4	1,650	600
5	2,300	650
6	3,000	700

Evaluating the Scheme

Given the importance of graded readers for the learning of English and the large number of readers available, it is astonishing that there is so little research to guide the construction of a graded readers scheme. What is less astonishing but equally depressing is that the research that has been done seems to have little influence on the construction of such schemes. In this respect the Penguin Readers are no better or worse than other series of graded readers.

Graded readers should aim at taking learners to the stage where they can move from the highest level of graded reading to unsimplified texts without too much of a step. Ideally the gap between Levels 5 and 6 of the Penguin Readers should be a little smaller than the gap between Level 6 and easy unsimplified reading. Hirsh and Nation (1992, p. 693) in a study of unsimplified texts written for younger readers recommended that there should be readers at the 2,600-word level and 5,000-word level. The Penguin Readers' 2,300- and 3,000-word levels are a very good move in this direction. This recommendation is based on the idea that for extensive reading learners should know 98-99% of the running words in a text. That is, there should be one or two unknown words in every 100 words, roughly one in every 5-10 lines of text.

This principle also relates to the gap between each level of the

1321

scheme. Table 1 shows that the gap between the levels of the Penguin Readers goes from 200 to 550 to 600 to 650 to 700 headwords. The gap becomes greater with each successive level. This is sensible because, if the vocabulary in the scheme is properly graded according to valid frequency counts, then at each successive level it will take more different headwords to cover the same percentage of running words. Ideally, at each level a learner should know about 98-99% of the words. When the learner moves to the next level after having met all or most of the headwords at that level, the learner should know around 98-99% of the words at that new level. That is, the gap between each level should be large enough to allow a reasonable number of unfamiliar words to come in, and yet small enough to still allow the learner to read for pleasure because the unfamiliar words only make up 1-2% of the running words. Wodinsky and Nation (1988) in a study of two graded readers at the 1,100-word level (Longman Level 4) found that the move from the previous 750-word level (Longman Level 3) (a gap of 350 words) gave 93.1% and 96.5% coverage of the new level. This was only a small study done before desktop computers were readily available, but it suggests that gaps of around 600 words between levels may be too big, especially between Levels 2 and 3, and 3 and 4. Clearly, research is needed and could very easily be done to empirically check the optimal gaps between the various levels of a graded reading scheme.

Research is also needed to see how many readers a learner would need to read at each level in order to (a) meet all of the words at that level and (b) have a good chance of learning most of the words at that level. Wodinsky and Nation's (1988) study showed that reading two texts at the 1,100-word level resulted in meeting only 57% of the unfamiliar words at that level. Similarly, reading only two readers provided little repetition of the unfamiliar words. Only 56 out of the 350 available words occurred 10 times or more in the two readers studied. If graded readers are to be used as a means of vocabulary expansion, it would be useful to know how many should be read at each level to get enough meetings with the vocabulary at that level. Given the computing technology and computer programs available, this is easy research to do.

About twenty words in each reader appear in bold type and are glossed at the back of the book. The selection of these words is a little strange. Some of them are topic related words probably beyond the vocabulary allowed at the level. The bolding and glossing draw attention to the words and increase the chances they will be learned. It would be more sensible to draw attention to words within the level rather than outside the level, as they may appear in the other readers at that level

322

and beyond.

The grading of grammatical structures in the Penguin Readers scheme, like virtually all other schemes and like most, but not all, course books, ignores the information on verb form frequency (George, 1963; Ota, 1963; Kramsky, 1969) and other corpus-based research (Kennedy, 1992) that has been around for decades. For example, present simple (future meaning) is in Level 5 although this is one of the most frequent ways of referring to the future. Present continuous is at Level 1 although for every one occurrence of present continuous there are at least 24 occurrences of simple present meaning "now". And so on.

The argument could be put that graded readers should reflect the grading schemes of course books because graded readers are supplementary materials. There are two answers to this argument. First, research on learning a second language through reading (Elley, 1991; Elley & Mangubhai, 1981) has shown that extensive reading can be a major contributor to second language proficiency. Texts for extensive reading should not distort a critical source of language input, graded readers, by using unsound structure grading schemes. Second, if course design is to take account of research, this has to be started somewhere. If course books will not do it, then the graded reader series should take the lead.

Fortunately, grading of grammatical structures is not a critical part in graded readers. Vocabulary grading is much more important, and it may be just as effective to set rough sentence length guidelines and leave the control of grammatical difficulty to the intuitive, sensitive judgement of writers and editors who have a feel for the learners who will read the books.

The *Penguin Readers* scheme, like other graded reader schemes, lacks a sensible base informed by research. This shortcoming does not damn it as a useful scheme but makes it apparent that yet another opportunity for substantial development in this critical area has not yet been taken.

Evaluating the Readers

The late Colin Mortimer, talking about materials development in English language teaching (ELT), made the point that when Shakespeare wrote his plays he was working under several constraints. Because the plays were performed live, you could not have flashbacks to past events. Thought could not be directly revealed, action was limited to what could be shown on a stage. And yet, Shakespeare created masterpieces. Writers of ELT materials and graded readers also work under

1323

constraints, particularly constraints brought about by the learners' proficiency level. And yet, within these vocabulary and grammar constraints, it should be possible to create little masterpieces. Hill and Thomas (1988, p. 46) make a similar point. The little masterpieces in the Penguin Readers series seem to come when the grammar controls are not so strictly applied. *Babe*, at Level 2, is a very enjoyable read. *Braveheart* at level 3 and Seven at level 4 are gripping stories. Hill (1997) recommends *Misery* (Level 6), *Surfer* (Level 1), and *Mrs Packletide's Tiger and Other Stories* (Level 6), but these were not available to the reviewer when this review was written.

The fact that many of the readers are also available as (unsimplified) films opens up interesting possibilities. With the film-book connections, it would seem to be a relatively easy job to get learners fired up about extensive reading. Some of the readers are a bit stilted because of the control and cutting, but the joy of being able to read a well known book in a second or foreign language, and understand it, will help overcome this.

The Penguin Readers are to be welcomed. The exciting range of titles, their attractive presentation, and their ready availability will undoubtedly make them a commercial and educational success.

References

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*(3), 375-411.

Elley, W. B., & Mangubhai, F. (1981). The long-term effects of a book flood on children's language growth. *Directions, 7*, 15-24.

George, H. V. (1963). A verb form frequency count. *ELT Journal, 18*(1), 31-37.

Hill, D. R. (1997). Graded (Basal) readers—choosing the best. *The Language Teacher, 21*(5), 21-26.

Hill, D. R. and Thomas, H. R. (1988). Survey review: Graded readers (Part 1). *English Language Teaching Journal, 42*(1), 44-52.

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2), 689-696.

Kennedy, G. (1992). Preferred ways of putting things with implications for language teaching. In J. Svartvik (Ed.) Directions in corpus linguistics. Berlin: Mouton de Gruyter. *Trends in Linguistics: Studies and Monographs, 65*, 335-373.

324

Kramsky, J. (1969). Verb form frequency in English. *Brno Studies in English*, *8*, 111-120.

Ota, A. (1963). *Tense and aspect in present-day American English.* Tokyo: Kenkyusha.

Wodinsky, M., & Nation, P. (1988). Learning from graded readers. *Reading in a Foreign Language, 5*(1), 155-161.

1325

Analysing Casual Conversation. (1997). By Suzanne Eggins and Diana Slade. London and Washington: Cassell.

Reviewed by MARTHA S. BEAN
San José State University

In *Analysing Casual Conversation*, Suzanne Eggins and Diana Slade undertake on the analysis of casual conversation, a genre with far more diffuse boundaries than other speech events and speech acts that have emerged in academic discourse. The authors distinguish casual conversation from other talk by defining it as "talk which is NOT motivated by any clear pragmatic purpose" and set out to establish a "comprehensive set of techniques for analysing patterns ... at a variety of linguistic levels" as well as to explore "how the micro patterns of difference realized in everyday chat enable meaning of relevance to the culture at large to be negotiated and contested." This attempt to find the nexus of microlevel and macrolevel concerns is an ambitious goal not fully realized, but the reader is grateful that the authors have tried as their work helps us to locate casual conversation as an important means, perhaps the primary means, for negotiating "social identity and interpersonal relations." The points made in chapter one, as in every other chapter, are amply illustrated through long segments of naturally occurring talk.

In Chapter Two, the authors compare and contrast the principal approaches to work on talk; namely, the ethnomethodological or conversational analytic (CA) approach from the work of Sacks, Schegloff, and Jefferson; the sociolinguistic approach from the work of Hymes, Gumperz, and Labov; the logico-philosophic approach from the work of Grice and Levinson and others; the structural-functional approach from the work of Sinclair and Coulthard; and the social-semiotic approach from the work of Halliday and Hasan. Adding yet another dimension to ways of viewing talk, the authors likewise cite work in critical linguistics (Kress and others) and critical discourse analysis (Fairclough and others), pointing out the potential of such disciplines to uncover hidden power relations in everyday exchanges. In their orientation towards such critical analyses of talk, the authors part company with linguists such as Tannen (1990) who take a more neutral stance on the differential effects of discourse. The authors here set up the analytical frame for the rest of the book namely the four types of patterns that jointly construct social as well as propositional meaning in talk; grammatical, semantic, discoursive, and generic.

In Chapter Three, the authors apply their fist level of analysis to

1326

examine the grammar of casual conversation, contending that the analysis of clause in terms of propositional structure and grammatical mood is crucial to revealing inevitable "tensions between equality and difference" (p. 67). Following Halliday (1994), the authors identify the declarative, imperative, wh-interrogative, polar (yes-no) interrogative, and exclamative moods as the major clause categories and invite readers to seek clues to speaker identity by examining mood in the talk of the three participants in the first focus conversation. The careful coding of the two-page segment of talk that follows allows readers to replicate the ensuing analysis. By charting mood across a conversation between middle class Australian parents and their adult son, the authors show how each speaker uses mood to create a social role within the context examined. The authors suggest that such unconscious patterns of clause-type choice achieve nothing less than the crafting of social identity, in this case along the somewhat anticipated lines of middle class white Australian family culture.

Taking a semantic stance in Chapter Four, the authors first reprise their analysis of mood in a conversation among five working-class male workmates, adding yet another analytical dimension by applying Martin's notions of appraisal and involvement to show how speakers signal varying degrees of solidarity and intimacy with their interlocutor (Martin, in press). In Martin's scheme, appraisal pertains to lexical choices that signal appreciation, affect, judgement, and amplification while involvement includes the use of naming, technicality, swearing, and slang. Again, patterns established for each speaker achieve different social roles. The authors find the primary goal of the workmates' talk to be evaluating and regulating group member behavior through appraisal and involvement. In addition, they foreground the seminal role of humor in mitigating and amplifying degrees of "otherness" and "in-ness."

Chapter Five applies Halliday's functional-semantic systemic approach to interaction at a five-person, mixed gender dinner party. The authors address the kinds of moves used in the dinner party exchanges according to function; namely, opening, reacting, developing, prolonging, and appending as well as supporitng and confronting moves, each of which is broken down into numerous subfunctions. Uncovering the kinds of moves used by individual speakers again allows the analysts to discern who is playing which role in the exchange, that is, whether speakers tend to initiate or respond in a given exchange and to confront or support within and what that means for both the speakers and the culture they are constructing through their choices. Perhaps surprisingly, the authors find that support

moves tend to terminate conversation while confrontation moves tend to prolong it.

In Chapters Six and Seven the authors take a more global view of conversation by moving from chat to conversational chunks with a more discernible structure, like stories and gossip. Because each can be described and recognized, Eggins and Slade argue for a generic approach to analyzing points on the conversational continuum ranging from chat to gossip to opinion to story-telling. They characterize the chat end as more interactive and less tightly structured and the story end as less interactive and more tightly structured, thus allowing analysts to focus on "macro text structures." The analysis of the talk about "Tamara" is particularly enlightening as it gives a clear demonstration of the construction of attitude as speakers verbally align themselves *with* certain people and situations and *against* others. On another note, although the authors avoid emphasizing gender, the analyses found in Chapters Four and Seven strongly suggest that men favor humor as a means of social control while women favor gossip (viewed positively here).

Analysing Casual Conversation falls squarely within the social constructionist view of language and society. The authors are to be congratulated for tackling longer segments of talk in very different social settings. The systematic application of four analytical perspectives to these segments is thoroughly and effectively executed, leaving readers with renewed appreciation for the infinite subtlety and nuance of casual conversation and how it is achieved through the legion of strategic grammatical, semantic, and discursive devices available to speakers. However, while Eggins and Slade show readers how to layer analysis upon analysis and select the best perspective for the particular segment under consideration, they are less successful in taking steps towards (a) creating a model that brings the different perspectives into alignment as a whole, and (b) despite nods to Kress and other critical discourse analysts, establishing more explicit connections with macro-level concerns, i.e., constructs of culture. For that, perhaps only Vygotsky (1978) will do. Still, the book has brought together "under one roof" so to speak, virtually all the major approaches to the analysis of conversation in a work that is sure to be foundational in the further study of casual talk.

References

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar* (2nd ed.). London: Edward Arnold.

Martin, J. R. (in press). *Beyond exchange: Appraisal systems in English.* In S. Hunstan and G. Thompson (Eds.), *Evaluation in text*. Oxford: Oxford University Press.

Tannen, D. (1990). *You just don't understand: Men and women in conversation.* New York: Ballantine Books.

Vygotsky, L. (1986). *Thought and language.* Cambridge: The Massachusetts Institute of Technology.

1329

German: a self-teaching guide. (1997). By Heimy Taylor and Werner Haas. 2nd ed.. New York: John Wiley & Sons. Pp. viii, 279, Soft cover $17.95. Audiocassette $9.95 (not available for review).

Reviewed by BRIGITTE OLSEN
Defense Language Institute Foreign Language Center

This textbook provides for the self-study of German, focusing primarily on "important communicative skills" (p.vii). With an Introduction to German sounds and spelling, followed by ten chapters of high-frequency topical material, the authors guide the self-learners through pronunciation and grammar rules, acquaint them with culture capsules, and supply communicative practice along with relevant exercises. At first glance, one notices a lack of visual appeal, customary in modern textbooks. There are neither pictures nor illustrations. Its content is, however, well presented and pleasantly arranged in clear print on the pages. The only colors used—black, red, and gold—have been reserved for the soft cover, suggestive of the German flag.

The individual lessons are well organized. For instance, Chapter One: Greetings, Introductions, and Useful Expressions begins with important Culture Notes relating to the chapter content, followed by *Dialoge* and *Kommunikation*. Here one finds also pertinent expressions, a listing of the important *wh*-words, later applied in the *Aktivitäten A, B, and C*. Next, the *Grammatik* explanations are clearly expressed and the exercises present copious, but not contextualized material for practice, rendering a workbook superfluous. *Sprechen Leicht Gemacht*, a further speaking activity coupled with translation exercises allows for personalization; *Wiederholung* provides more practice to which answers are given at the end of the book. The *Wortschatz* is broken down into a useful listing of nouns, verbs, idiomatic expressions, and cognates. A self-learner should be pleased with himself and with the text, being able to obtain so much knowledge in just one lesson.

The subsequent nine chapters follow the effective format with topics in logical progression. Chapter Two not only contains numbers, but numbers in context with measurements, currencies, sizes, and temperatures, both in the American measure and the metric system. The topics of the next chapters focus on food, lodging, and telling time. Chapter Six deals with health care. Unlike other texts at the beginner's level, it provides students in the Culture Notes with essential terminology such as *Notruf, Sonderklasse and Nachtapotheke*. The final lessons focus on sports, communication at a number of public places like the airport, and concludes with autobiographical data. Though the

1330

authors comment that this book "lays no claim to completeness" (p.vii), I believe the self-learner is exposed to a wealth of German, applicable to the basic level of communication so as to be "able to satisfy minimum courtesy requirements and maintain very simple face-to-face conversations on familiar topics" (ILR Skill Level Description—Level 1 Speaking).

Nevertheless, minor inaccuracies or oversights in the text could somewhat hamper the self-instructional process because the self-learner by the very nature of learning by himself is on his own. For instance, while the instructions in Chapter Six, *Wiederholung*, are clear "Write a question replacing the boldface word(s)" (p.143), the example is not:

BEISPIEL Jutta hat hohes Fieber.
Wer hat hohes Fieber?
(Jutta hat hohes Fieber.)

Neither is sentence 5 printed correctly. *Letzte Woche fuhr Herr Schwarz nach Köln*. Somehow, Köln missed its boldface.

Furthermore, in Chapter One, *du* and *ihr*, the familiar address forms, are listed under the three subject pronouns and their use is properly explained. Yet the practice of *du* occurs rather sparingly, only once in most exercises and practice sections; *ihr* has no application at all (p.16). Anyone who has traveled to Germany lately realizes that the familiar forms are being used more frequently than before and, consequently, more practice seems desirable. Along the same lines, the absence of *du* and *ihr* in the *Grammatik* section of Chapter Three relating to command forms appears problematic.

Also in Chapter One, a need to explain contrastive grammar becomes apparent. The three forms of present tense in English were pointed out and set against the single form in German (p. 19). However, *Wie lange sind Sie schon in Deutschland*? (present tense) in the *Dialoge* is merely translated to "How long have you been in Germany?" (present perfect, p. 12). Without an explanation, this presumed incongruence might give a self-learner cause for concern.

Another grammar difficulty rests with the simple past tense. The explanations accurately define the simple past tense (preterit, p. 133). Yet the usage in the *Dialoge* does not seem authentic. It is unlikely that a German Elke would answer *Ich nahm vor einer Stunde wieder zwei [Tabletten]*, or *Ich trank seit gestern bestimmt zwei Liter Mineralwasser*. As the preterite represents the grammar focal point of Chapter Seven, a good opportunity was missed here by not including authentic reading texts allowing for proper use of the preterit, espe-

1331

cially since the authors' intention was also to provide "basic tools to... read... simple German" (p.viii).

While for the most part, the *dialoge* are in convincing German, the very first one (p.12) gives pause to think. Herr B. and Herr C. meet for the first time: After the exchange of names, Herr C. then asks immediately, *Wie geht es Ihnen?* Although an American can ask "How are you?", German custom reserves that question for acquaintances, because it means more than just hello.

A final point about the dialogue section. The instruction given, "Read the dialogues aloud in German and check the English translation," does not state a purpose. Check for what? To learn that there is no word-for-word translation? To check for authenticity in English? If so, a student would perhaps find a glitch in Chapter Six: *Noch diese Woche* does not translate into "This coming week" (p. 129). In Chapter Seven the following dialog appears to lack authenticity:

- What's on TV tonight?
- Nothing special. On channel 1 they're showing a well-known detective film, on channel 2 a silly entertainment show, and on channel 3 a boring documentary about Brazil.

- What should we watch?
- Nothing. I would rather read the new *Spiegel* and listen to something nice on the radio (S. 86).

Even though one would wish for more perfection when a text is written for self-study in particular, *german: a self-teaching guide* still is well conceived and developed. Certainly I concur with the authors that, "A *self-teaching guide* will not turn you into a polished speaker of German overnight... It opens the gate to a very gratifying experience: understanding and appreciating the language and culture of the German-speaking people." To that end, the book will suffice and the learner will succeed.

1332

Colloquial Vietnamese: A Complete Language Course. (1994). By Tuan Duc Vuong and John Moore. London & New York: Routledge.

Reviewed by NGUYEN DINH-HOA
Southern Illinois University

This is another textbook in the Routledge "Colloquial" Series, designed for English speakers. It consists of seventeen lessons built around seventeen topical chapters, each consisting of a dialogue, exercises, and grammar notes aiming at teaching the daily speech of Vietnam. In addition to the key to the exercises (pp. 249-287), *Colloquial Vietnamese* includes a summary on grammar (pp. 224-248), a Vietnamese-English glossary (pp. 288-302), and an English-Vietnamese glossary (pp. 303-314).

This well-organized book addresses adult college students in evening classes as well as tourists and businesspeople. The dialogues are well constructed, and the exercises should be helpful for someone who wants to learn both the spoken language and some written Vietnamese, with the assistance of repetition, match-up, substitution and translation drills.

Being the fruit of the collaboration between a native speaker (the first author) and an English diplomat (the second author), this manual is a useful learning tool. It is regrettable, however, that for some reason the authors have allowed a large number of mistakes to creep into their text, especially among the cultural notes.

The Introduction, which follows a map of Vietnam and deals with the population, history, economy, culture, and language, is an example. The sentence on the Lac (p. 2) contains some erroneous information: These people are said to inhabit "the Red River Delta and the central regions and the Muong." The latter name refers to the Möôøng minority "nationality," considered racial and cultural sisters of the Vietnamese. The present country name Viet Nam was not "received in 1945," but had been used earlier, even during the French period. The "August Revolution" started with the uprising in Hanoi on 19 August, not on 16 August, 1945.

Some definitions and explanations leave much to be desired: "Chöõ means word and noâm means prose which is easy to understand" (p. 5). Actually chöõ means 'letter; script, character', and *noâm*— derived from *nam* 'south'— is used to refer to the "southern script," i.e., the demotic system of writing, which borrowed Chinese characters to transcribe native words. On the other hand, the Roman script called

1333

quoác ngöõ was not "introduced" by Alexandre de Rhodes, nor has it "replaced the Chinese calligraphic script officially" "since then." The Jesuit missionary-scholar from Avignon was rather the codifier of that writing system, which had been created by Portuguese missionaries, two of whom de Rhodes mentioned in the preface of his Vietnamese-Portuguese-Latin dictionary, published in Rome in 1651.

P. 6: there are not "a few cases" of polysyllabism, but many two-syllable, and even three-syllable and four-syllable words—through reduplication and compounding—and they have not been "made by hyphenation," either. Indeed, *sô-mi*, a loanword from French, is not a compound word, but simply a disyllabic borrowing—unlike the example cited later, *hoan-ngheânh*—a Chinese loanword—which is not two words, but a single word with two syllables! The combination of two vowel letters in the example ao (p. 7) does not represent "one distinct sound," but rather the vowel /a/ followed by the semivowel /-w/.

In the discussion of the sound system (Chapter 1), the authors show confusion between the sounds and their transcriptions. It is indeed misleading to state on p. 10 that vowels number 6 and consonants number 16, then on pages 10-11 to list as many as 24 consonants and only 6 "main vowels"(!). The truth is that the nucleus of each syllable in its canonical form has either one of the *eleven* simple vowels or one of *three* complex vowels */ia, ua, öa/*. The confusion is also due to mixing vowels with semivowels, and again to mixing sounds and letters representing them. Typically the initial consonant of *tha* is not /O/ (as in English *thin*), but simply an aspirated [t'] (as in English *tin*).

Errors in translation and interpretation can be easily spotted. The verb 'to live' in English should not always be translated as soáng because in the context, Toâi soáng ôû Luaân Ñoân, it means 'to reside, dwell'.

P. 55: 'fourteen' cannot be *möôøi tö*.

P. 65: *maùy laïnh* means 'air conditioner', not 'air conditioning'

Pp. 67, 226: nhöõng does not mean a larger number. *Caùc* is just a pluralizer borrowed from Chinese, whereas the native word *nhöõng* is used when the noun is further modified (by a relative clause, for instance): 'those (whom) I met yesterday' would be *Nhöõng ngöôøi (maø) toâi gaëp hoâm qua*. Also *caùc* is used before status pronouns in address, as in *Xin caùc oâng, caùc baø giöõ choã tröôùc* 'Please, ladies and gentlemen, make reservations in advance' (cf. *nhöõng oâng baø [naøo] chöa giöõ choã* 'those of you who haven't made reservations').

'Better than' in X speaks French better than English should be *gioûi hôn, khaù hôn*, not *toát hôn* (pp. 68, 83, 87, 90).

P. 89: since *quâc* already means 'country', Vietnamese do not say *nöôùc Trung Quoác*. People usually say, *Trung-quâc, Trung-hoa*—or *nöôùc Taøu*.

Pp. 102, 104, 108: The interrogative sentence *OÂng muoán baûo chi vaäy?* is awkward and reminds me of a lecturer in 1975 who talked and wrote a kind of Vietnamese that Vietnamese students and diplomats in London at that time found archaic and funny. Actually it does not mean 'What do you want to ask?' but rather 'What kind of advice and counsel do you want to give me, Sir?' OÂng caàn gì? would be much more natural, and less subservient- or rural-sounding.

The polite expression *Caùm ôn*, which means only 'to thank', should be followed by the appropriate second-person pronoun. The same applies to Xin loãi... 'Excuse me'. Used alone, these expressions are translationese and not polite!

P. 168: The word ñeå introduced in Lesson 14, can serve both as a preposition 'in order to' and a conjunction 'in order that [something happens]. One does not start a sentence with a prepositional phrase like *ñeåñi buoân*.

Pp. 169, 172, 179: *nhöõng ngöôøi baïn cuõ Vieät Nam cuûa toâi* 'my old Vietnamese friends' sounds awkward. The order *nhöõng ngöôøi baïn Vieät Nam cuõ cuûa toâi* is better. But a native speaker would probably say *baïn cuõ Vieät Nam*, omitting both the plural form and the possessive form: Not everything in the source language has to be translated into something "equivalent" in the target language.

Pp. 188-89: *taát nieân* does not mean 'New Year', but 'year-end'.

P. 196: The idiom *huoáng chi (= huoáng gì, huoáng hoà)* does not mean 'more or less', but 'all the more reason for ... ' if the preceding affirmative clause contains coøn—or 'much less, let alone' if the preceding interrogative clause contains *coøn*. Example: *Ngöôøi döng anh ta coøn giuùp ñöõ, huoáng chi choã beø baïn!* He helps even strangers, all the more reason for him to help his own friends'. (**Töï ñieån tieáng Vieät, by Hoaøng Kheâ**., Editor, et al. 1994).

Somehow the book overlooks the very important difference between *chuùng ta* 'inclusive we' and *chuùng toái* 'exclusive we'. There are scores of errors in spelling and in placing tone marks, but these typos can easily be corrected by a rigorous editor. Despite the infelicities enumerated above, this well illustrated handbook is a welcome addition to the Vietnamese language library shelf, and certainly a serviceable quick learning tool.

1335

General Information

Calendar of Events*

1998

5-7 March, *Ohio Foreign Language Association*, Cincinnati. Information Ann Rolston, (614) 385-6317, Email [annfrench@hocking.net]. American Association of Applied Linguistics, 14-17 March, Seattle. Information AAAL, PO Box 21686, Eagan, MN 55121-0686; (612) 953-0805, Fax (612) 431-8404, Email [aaaloffice@aaal.org].

17-21 March, *Teachers of English to Speakers of Other Languages, Seattle*. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703) 836-0774, Fax (703) 836-7864, Email [conv@tesol.edu].

26-28 March, *Balkan and South Slavic Languages, Literatures, and Folklore*, Tucson. Information John Leafgren, Department of Russian and Slavic Languages, 340 Modern Languages, University of Arizona, Tucson, AZ 85721; email [leafgren@ccit.arizona.edu].

26-29 March, *Central States Conference on the Teaching of Foreign Languages*, Milwaukee. Information CSCTFL, Rosalie Cheatham, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; (501) 569-8159, Fax (501) 569-3220, Email [rmcheatham@ualr.edu]. Society, Language, and Culture in Post-Communist Russia, the Other Former

2-4 April, *Republics of the Soviet Union, and Eastern Europe*, Lubbock. Information Lewis Tracy, Classical & Modern Languages & Literatures, Texas Tech University, Lubbock, TX 70409-2071; Email [pvlew@ttacs.ttu.edu].

16-19 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu], URL [http://www.dickinson.edu/~neconf/].

23-25 April, *Southwest Conference on Language Teaching and Arizona Language Association*, Mesa. Information Carl H. Johnson, Texas Education Agency, 1701 N. Congress Avenue, Austin, TX 78701-1494.

1336

*Courtesy of *The Modern Language Journal* (University of Winsconsin)

24-26 April, *Pacific Northwest Council for Languages*, Boise. Information PNCFL, PO Box 4649, Portland, OR 97208-4649. 7-9 June, *National Association of Professors of Hebrew*, New York. Information NAPH Conference, 1346 Van Hise Hall, 1220 Linden Drive, University of Wisconsin, Madison, WI 53706; (608) 262-3204, Email [morahg@lss.wisc.edu].

16-20 July, *Society of the History of Authorship, Reading and Publishing*. Vancouver. Information SHARP 98, c/o Canadian Centre for Studies in Publishing, Simon Fraser University at Harbour Centre, 515 West Hasgings Street, Vancouver, BC V6B 5K3 Canada; (604) 291-5093, Fax (604) 291-5098, Email [dkirby@sfu.ca].

23-26 July, *American Association of Teachers of French*, Montreal. Information AATF, Mailcode 4510, Department of Foreign Languages, Southern Illinois University, Carbondale, IL 62901-4510; (618) 453-5731, Fax (618) 453-5733, Email [abrate@siu.edu].

31 July-4 August, *American Association of Teachers of Spanish and Portuguese*, Madrid. Information AATSP, Frasier Hall #8, University of Northern Colorado, Greeley, CO 80639; (970) 351-1090, Fax (970) 351-1095, Email [seelsandste@bentley.univnorthco.edu].

24-27 September, *American Association for the Advancement of Slavic Studies*, Boca Raton. Information AAASS; Email [walker@core-mail.fas.harvard.edu].

September dates to be announced, *International Congress of Slavists*, Cracow. Information Robert Rothstein, Department of Slavic Languages, Herter Hall, University of Massachusetts, Amherst, MA 01003-3940.

37-31 October, *Foreign Language Association of North Carolina*, location to be announced. Information Mary Lynn Redmond, 6 Sun Oak Court, Greensboro, NC 27410; Fax (910) 759-4591, Email [redmond@wfu.edu].

20-22 November, *American Council on the Teaching of Foreign Languages*, Chicago. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

20-22 November, *American Association of Teachers of German*, Chicago. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

1337

27-30 December, *Modern Language Association of America*, San Francisco. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, San Francisco. Information Masako Ueda, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; (401) 863-3933, Fax (401) 863-7330, Email [masako_ueda@brown.edu].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, San Francisco. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

1999

3-9 March, *American Association of Applied Linguistics*, New York. Information AAAL, PO Box 21686, Eagan, MN 55121-0686; (612) 953-0805, Fax (612) 431-8404, Email [aaaloffice@aaal.org].

8-14 March, *Teachers of English to Speakers of Other Languages*, New York. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703) 836-0774, Fax (703) 836-7864, Email [conv@tesol.edu].

11-13 March, *Southern Conference on Language Teaching*, Virginia Beach. Information Lee Bradley, Valdosta State University, Valdosta, GA 31698; (912) 333-7358, Fax (912) 333-7389, Email [lbradley@grits.valdosta.peachnet.edu], URL [http://www.valdosta.edu/scolt/].

7-10 April, *Pacific Northwest Council for Languages*, Tacoma. Information PNCFL, PO Box 4649, Portland, OR 97208-4649.

8-11 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu], URL [http://www.dickinson.edu/~neconf/].

15-18 April, *Central States Conference on the Teaching of Foreign Languages*, Little Rock. Information CSCTFL, Rosalie Cheatham, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; (501) 569-8159, Fax (501) 569-3220, Email [rmcheatham@ualr.edu].

1338

July dates to be announced, *International Assocation of Teachers of Russian Language and Literature*, Slovakia. Information ACTR, 1776 Massachusetts Avenue NW, Washington, DC 20036; Email [ddavidson@actr.org].

18-21 November, *American Association for the Advancement of Slavic Studies*, St. Louis. Information AAASS; Email [walker@core-mail.fas.harvard.edu].

19-21 November, *American Council on the Teaching of Foreign Languages*, Dallas. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

19-21 November, *American Association of Teachers of German*, Dallas. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America*, location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, location to be announced. Information Masako Ueda, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; (401) 863-3933, Fax (401) 863-7330, Email [masako_ueda@brown.edu].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

2000

19-21 March, *American Association of Applied Linguistics*, Vancouver. Information AAAL, PO Box 21686, Eagan, MN 55121-0686; (612) 953-0805, Fax (612) 431-8404, Email [aaaloffice@aaal.org].

21-25 March, *Teachers of English to Speakers of Other Languages*, Vancouver. Information TESOL, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; (703) 836-0774, Fax (703) 836-7864, Email [conv@tesol.edu].

1339

13-16 April, *Northeast Conference on the Teaching of Foreign Languages*, Washington DC. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu], URL [http://www.dickinson.edu/~neconf/].

9-12 November, *American Association for the Advancement of Slavic Studies*, Denver. Information AAASS; Email [walker@core-mail.fas.harvard.edu].

17-19 November, *American Council on the Teaching of Foreign Languages*, Boston. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

17-19 November, *American Association of Teachers of German*, Boston. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America*, location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, location to be announced. Information Masako Ueda, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; (401) 863-3933, Fax (401) 863-7330, Email [masako_ueda@brown.edu].

27-30 December, American Association of Teachers of Slavic & E. European Languages, location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520)885-2663, Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

2001

19-22 April, Northeast Conference on the Teaching of Foreign Languages, New York. Information Northeast Conference, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; (717) 245-1977, Fax (717) 245-1976, Email [neconf@dickinson.edu], URL [http://www.dickinson.edu/~neconf/].

16-18 December, American Council on the Teaching of Foreign Languages, Washington, D.C. Information ACTFL, 6 Executive Plaza, Yonkers, NY 10701-6801; (914) 963-8830, Fax (914) 963-1275, URL [http://www.infi.net/~actfl].

340

16-18 November, American Association of Teachers of German, Washington, D.C. Information AATG, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; (609) 795-5553, Fax (609) 795-9398, Email [73740.3231@compuserve.com].

27-30 December, Modern Language Association of America, location to be announced. Information MLA, 10 Astor Place, New York, NY 10003-6981; Fax (212) 477-9863, Email [convention@mla.org].

27-30 December, North American Association of Teachers of Czech, location to be announced. Information Masako Ueda, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; (401) 863-3933, Fax (401) 863-7330, Email [masako_ueda@brown.edu].

27-30 December, American Association of Teachers of Slavic & E. European Languages, location to be announced. Information AATSEEL, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Fax (520) 885-2663, Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

1341

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (*ALL*) is to increase and promote professional communication within the Defense Language Program and academic communities on adult language learning for functional purposes.

Submission of Manuscripts

The Editor encourages the submission of research and review manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Research Article

Divide your manuscript into the following sections:

- Abstract
- Introduction
- Method
- Results
- Discussion
- Conclusion
- Appendices
- Notes
- References
- Acknowledgements
- Author

Abstract
Identify the purpose of the article, provide an overview of the content, and suggest findings in an abstract of not more than 200 words.

Introduction
In a few paragraphs state the purpose of the study and relate it to the hypothesis and the experimental design. Point out the theoretical implications of the study and relate them to previous work in the area.

1342

Next, under the subsection *Literature Review,* discuss work that had a direct impact on your study. Cite only research pertinent to a specific issue and avoid references with only tangential or general significance. Emphasize pertinent findings and relevant methodological issues. Provide the logical continuity between previous and present work. Whenever appropriate, treat controversial issues fairly. You may state that certain studies support one conclusion and others challenge or contradict it.

Method

Describe how you conducted the study. Give a brief synopsis of the method. Next develop the subsections pertaining to the *participants,* the *materials,* and the *procedure.*

Participants. Identify the number and type of participants. Specify how they were selected and how many participated in each experiment. Provide major demographic characteristics such as age, sex, geographic location, and institutional affiliation.

Identify the number of experiment dropouts and the reasons they did not continue.

Apparatus. Describe briefly the materials used and their function in the experiment.

Procedure. Describe each step in the conduct of the research. Include the instructions to the participants, the formation of the groups, and the specific experimental manipulations.

Results

First state the results. Next describe them in sufficient detail to justify the findings. Mention all relevant results, including those that run counter to the hypothesis.

Tables and figures. Prepare tables to present exact values. Use tables sparingly. Sometimes you can present data more efficiently in a few sentences than in a table. Avoid developing tables for information already presented in other places. Prepare figures to illustrate key interactions, major interdependencies, and general comparisons. Indicate to the reader what to look for in tables and figures.

Discussion

Express your support or nonsupport for the original hypothesis. Next examine, interpret, and qualify the results and draw inferences from them. Do not repeat old statements: Create new statements that further contribute to your position and to readers understanding of it.

Conclusion

Succinctly describe the contribution of the study to the field. State how it has helped to

resolve the original problem. Identify conclusions and theoretical implications that can be drawn from your study.

Appendices
Place detailed information (for example, a table, lists of words, or a sample of a questionnaire) that would be distracting to read in the main body of the article in the appendices.

Notes
Use them for substantive information only, and number them serially throughout the manuscript. They all should be listed on a separate page entitled *Notes*.

References
Submit on a separate page of the manuscript a list of references with the centered heading: *References*. Arrange the entries alphabetically by surname of authors. Review the format for bibliographic entries of references in the following sample:

> Dulay, H., & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 16* (1), 93-95.
>
> Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

List all works cited in the manuscripts in *References*, and conversely, cite all works included in *References* in the manuscript. Include in reference citations in the text of the manuscript the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that you are quoting originally appeared, e.g., (Jones, 1982, pp. 235-238).

Acknowledgments
Identify colleagues who contributed to the study and assisted you in the writing process.

Author
Type the title of the article and the author's name on a separate page to ensure anonymity in the review process. Prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

> JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Review Article

It should describe, discuss, and evaluate several publications that fall into a topical

category in foreign language education. The relative significance of the publications in the context of teaching realms should be pointed out. A review article should be 15 to 20 double-spaced pages.

Review

Submit reviews of textbooks, scholarly works on foreign language education, dictionaries, tests, computer software, video tapes, and other non-print materials. Point out both positive and negative aspects of the work(s) being considered. In the three to five double-spaced pages of the manuscript, give a clear but brief statement of the work's content and a critical assessment of its contribution to the profession. Keep quotations short. Do not send reviews that are merely descriptive.

Manuscripts are accepted for consideration with the understanding that they are original material and are not being considered for publication elsewhere.

Specifications for Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-AP-AJ
ATTN: Editor (Dr. L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run from 10 to 30 pages.

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784.

Review Process

Manuscripts will be acknowledged by the editor upon receipt and subsequently sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

Specifications for Floppy Disks

Preferably use Windows-based software. Format manuscripts produced on one of the DOS-based or Macintosh systems, as an ASQII file at double density, if possible. Please name the software used.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue. Requests for permission to reprint should be addressed to the Editor and should include author's permission.

1346

Applied Language Learning Comment Card

We are always interested in hearing from our readers. Please use this card to comment on this issue: _____.
Volume Number Date

Tear Out and Remove

_____

_____

_____

_____

_____

_____

_____

MARK ONE
- ☐ New Subscription
- ☐ Address Change

PLEASE PRINT

_____
Rank/Title First Name Middle Initial Last Name

_____
Institution Division Department

_____
Street City State/Country Zip Code

Tear Out and Remove

From:

_____

_____

_____

Official Business

Place
Stamp
Here

1348

Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5506
USA

1348

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
PRESIDIO OF MONTEREY, CA 93944-5006

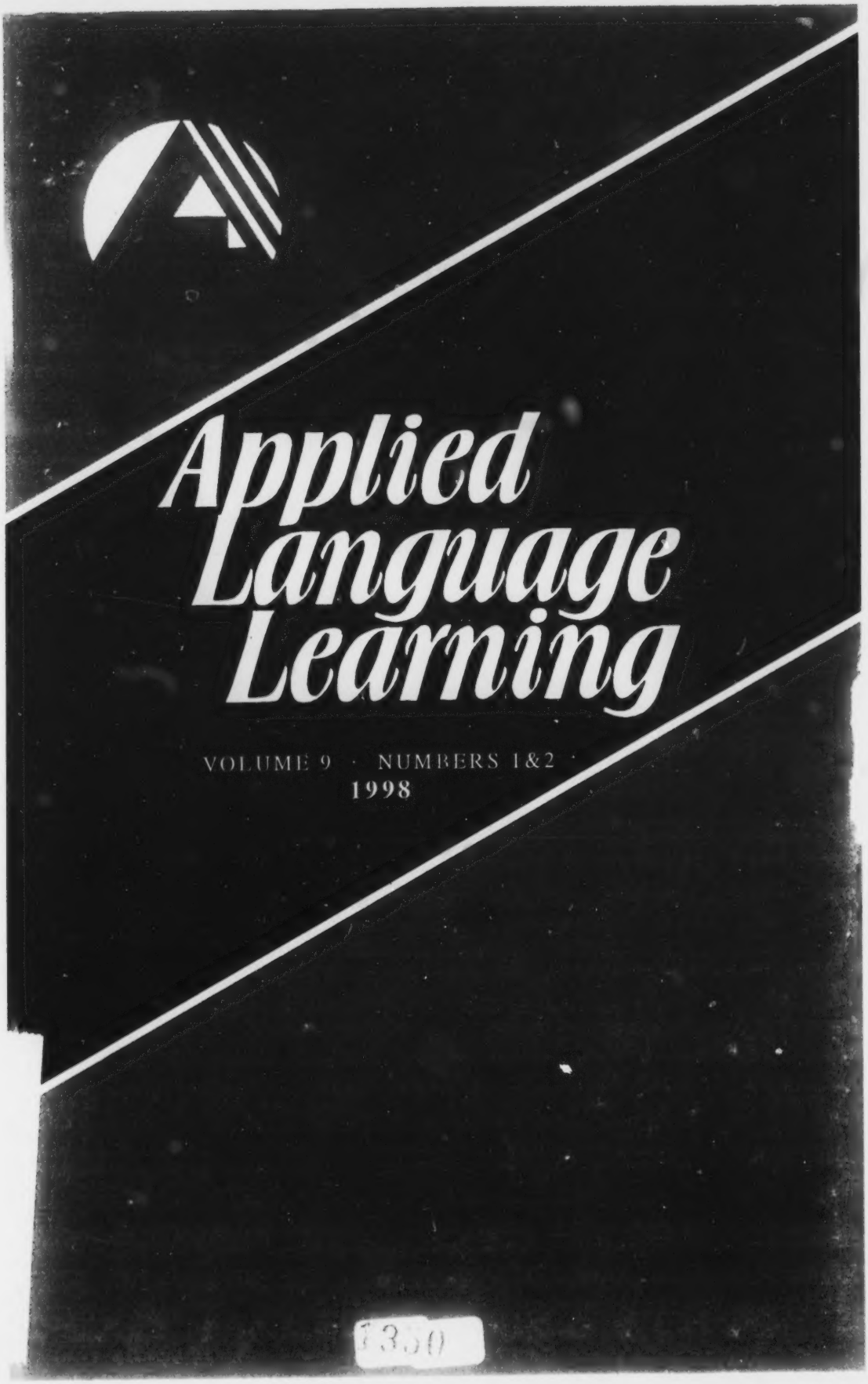
BULK RATE
U.S. POSTAGE
PAID
MONTEREY, CA
PERMIT NO. 42

PB-65-97-2
United States Army
PIN: 076 000-000
Approved for public release.
Distribution is unlimited.

1349

BEST COPY AVAILABLE

1349



Applied Language Learning

VOLUME 9 · NUMBERS 1&2
1998

350

Applied Language Learning

Professional Bulletin 65-98-1 & 2

1998 **Volume 9** **Numbers 1 & 2**

Lidia Woytak
Editor

Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006

1351

In Memory of

Richard A. Woytak

who always has believed that

Applied Language Learning

Will Create a Better World for All of Us

135

Applied Language Learning

PB 65-98-1 & 2

The mission of Professional Bulletin 65-98, *Applied Language Learning (US* ISSN 1041-6791), is to provide a forum for the exchange of ideas and information on instructional methods and techniques, curriculum and materials development, assessment of needs within the profession, testing and evaluation, and implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences.

Applied Language Learning, published semiannually by the Defense Language Institute Foreign Language Center and Presidio of Monterey, presents professional information. The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

By Order of the Secretary of the Army:

DENNIS J. REIMER
General, United States Army
Chief of Staff

JOEL B. HUDSON
Administrative Assistant to the
Secretary of the Army
05596

DANIEL D. DEVLIN
Colonel, United States Army
Commander
DLIFLC and POM

1353

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the authors.

Availability

To access *Applied Language Learning* on the Internet type:

http://lingnet.army.mil

Additionally, you may obtain the journal on microfilm from ERIC Clearinghouse on Language and Linguistics, Center for Applied Linguistics, 1118 22nd Street, NW, Washington, DC 20037.

Bulk rate postage is paid at DLIFLC. The basis of official distribution is one copy per training instructor and one per five military linguists.

Postmaster

Send change-of-address information to:

Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5006

United Parcel Customers

Location is *Applied Language Learning*
Tin Barn
Bldg. 518, Room 7
Presidio of Monterey, CA 93944-5006

Readers

Contact Editor, Dr. Woytak, by:

E-mail: woytakl@pom-emh1.army.mil
Telephone: (831) 242-5638
DSN: 878-5638
Fax: (831) 242-5850

Wordprocessing and graphics — *Specialist Richard S. Slone*, *Private Timothy S. Hughes*

Cover design — *Barney Inada*

1354

Reviewers for *Applied Language Learning*

The individuals listed below served as reviewers of manuscripts submitted to *Applied Language Learning* in 1998. We express our gratitude for expert service to:

Martha S. Bean
San José State University
Christine M. Campbell
Defense Language Institute Foreign Language Center
John B. Carroll
University of North Carolina
Marianne Celce-Murcia
University of California, Los Angeles
Dan Douglas
Iowa State University
Roderic A. Gale
Defense Language Institute Foreign Language Center
Robert C. Gardner
University of Western Ontario
Rafael Gomez
Monterey Institute of International Studies
John S. Hedgcock
Monterey Institute of International Studies
Evelyn Hatch
University of California, Los Angeles
Eli Hinkel
Seattle University
Gordon Jackson
Defense Language Institute Foreign Language Center
Renee Jourdenais
Monterey Institute of International Studies
Dorry M. Kenyon
Center for Applied Linguistics
Stephen D. Krashen
University of Southern California
Masaki Kobayashi
University of British Coumbia
James F. Lee
University of Indiana
Virginia LoCastro
International Christian University
Barbara Matthies
Iowa State University
Naoko Matsuo
Monterey Institute of International Studies
Denise Murray
San Jose State University
Paul Nation
Victoria University of Wellington
Rebecca Oxford
University of Alabama
Richard Schmidt
University of Hawaii-Manoa
David J. Shook
Georgia Institute of Technology
Roberta J. Vann
Iowa State University
Leo Van Lier
Monterey Institute of International Studies
Dolly J. Young
University of Tennessee

Applied Language Learning

1998 **Volume 9** **Numbers 1 & 2**

Articles

1 Language Aptitude Testing:
Learners and Applications
James R. Child

11 Zero-Based Language Aptitude Test Design:
Where's the Focus for the Test?
Pardee Lowe, Jr.

31 The Modern Language Aptitude Test for Predicting Learning
Success and Advising Students
Madeline Ehrman

71 Factors in the Prediction of Achievement and Proficiency
in a Foreign Language
Richard Sparks, James Javorsky, Jon Patton, and Leonore Ganschow

107 Essay Scores as Instruments for Placement and
Advancement in an Intensive English Program
Lynne Davis, Ruth Johnson, and Floyd Olive

News and Views

121 Interpreter in Action: Interview with Lieutenant Colonel
Richard Francona (Retired)
Lidia Woytak

135C

Reviews

147 Cots: Teaching by Chatting.....Leo Van Lier
149 Kenny and Savage (Eds.): Language and Development: Teachers in a Changing World.....Kevin W. K. Chu
151 Lee and VanPatten: Making Communicative Language Teaching Happen: Directions for Language Learning and Teaching.....Phillip A. White
154 Gates: The Road Ahead.....Roderic A. Gale

General Information

157 Calendar of Events
163 Information for Contributors

1357

Applied Language Learning
1998, Vol. 9 Nos. 1 & 2, pp. 1-10

Language Aptitude Testing

Learners and Applications

James R. Child
National Cryptologic School

The present article reviews language aptitude testing dilemmas both from the perspective of the National Security Agency (NSA) and of all the government organizations with language missions. First, it considers the possible need to cross-train linguists in government employ, sometimes from "difficult" into "easier" languages, but more frequently in the other direction. In so doing, it recommends which of three available aptitude measures is most suitable. Second, it takes up the relationship between language aptitude and the Interagency Language Roundtable (ILR) skill levels: Which test is appropriate for what skill? Can an aptitude test by itself or together with other measures predict learning success at Level 3 and up? Finally, the article treats the varying degrees of distance between English and other (selected) languages with reference to the aptitude model best suited to the purpose.

The need has become increasingly acute in U.S. society for persons with the four-skills described in the Interagency Language Roundtable (ILR) statements: Speaking (S), Listening (L), Reading (R) and Writing (W). It would seem arguable that any society would be enriched when a significant number of its members can attain at least a Level 2 competence in a given language in one or more skills rated over a range of six levels. Whether the enrichment is a matter of extending personal horizons for social and cultural purposes, or developing skills for the workplace, is less immediately important than the learning attainment per se.

1358

Candidates for Aptitude Testing

The point of departure in considering NSA's requirements and responses to them is the assumption that the language workforce is made up by and large of native speakers of English who have formally learned or otherwise acquired one or more foreign languages. Naturally those persons already skilled in languages the agency requires will likely spend significant portions of their careers processing texts in those languages, at the same time attempting to improve their English writing skills, in order to present better what they have understood, in translation, summary or other form.

However, as changing conditions dictate, it does happen on occasion that linguists in the hiring pipeline or already employed need to be cross-trained to other languages. This problem, fairly widespread in the "world of work" in general, regularly surfaces at NSA. Consideration is given below to the dimensions of the problem and the ways in which it can be handled effectively.

On-Board Working Linguists

Persons who have been hired as linguists and spent sufficient time doing solid professional work in second (or, less frequently, second and third) languages have shown themselves to be excellent candidates for the learning of other, usually "exotic" tongues. The kinds of language processing typical of the "world of work" at many government agencies include translation or interpretation as "top-down" skills, with subsidiary requirements to extract from (whole) products of those skills the critical information needed, in the form of summaries, paraphrases and the like. These activities may be viewed in the government frame of reference as performance measures. That is to say, they are carried out according to the letter and spirit of mission requirements which by their nature demand familiarity and experience with a wide array of content areas (or, "real-world knowledge"). Linguists with a "track record" in coping with one or several of these areas in language X have shown that they can bring much of that knowledge and experience to language Y, with the caveat that Y share a fair amount of basic structural features with X. Absent those commonalities, no amount of subject matter familiarity can in itself predict success in mastering a difficult language. Therefore NSA has opted to administer two aptitude tests—one, the Defense Language Aptitude Battery (DLAB) developed by Petersen and Al-Haik in 1973 as a measure to indicate likelihood of success in learning languages "structurally close" to English; the second, a test James Child developed at

NSA in 1973 called VORD (not an acronym) as an instrument for placing persons in classes in which the language is vastly different in structure from English. The notion of "distance" between and among languages in terms of difficulty will be discussed at a later point; for the moment it is enough that the two tests are administered to employees in conjunction with other measures relevant to cross-training, and that (at least in the case of DLAB) aptitude testing may simply be supportive of other indicators.

Prospective Linguists

Persons in this category normally have one thing in common: little experience in doing the kinds of language work many agencies require. Thus they must spend considerable time in becoming familiar with various specialized topic domains even when they would be using the languages which they bring with them. Without the transferable skills of on-board persons, aptitude testing might prove a critical element in helping managers and senior level technical experts make sound cross-training decisions in placing new hires in language training for which they were not originally programmed.

Language Uses: Skills and Levels

The three skills in demand in agencies taken as a whole are speaking, reading, and listening. (A fourth, writing, has always been somewhat marginal for the government). They are covered for government purposes, in the ILR definitions, in a six-tier system, from Level 0 (extremely limited memorized skills) to Level 5 (that of a highly educated native), with "plus-levels" in between each base level.

Of the skills, reading and listening (the receptive skills) are critical for those agencies which do not have major missions requiring interaction with speakers of (particular) languages, while speaking must obviously be added when such interaction is called for. As to the levels of attainment most frequently in demand in the world of work, it is fair to say that Levels 2 and 3 come into play most often in day-to-day operations, with a Level 4 competence occasionally required.

What does all of this imply for language aptitude testing? First, Level 2 skills should be targeted first, for Level 1 language is generally reduced in content and form to phrase- and sentence-length units containing material usually of little intrinsic interest to serious study: greetings; weather information; arrivals and departures of carriers and the like; all amenable to brute memory. Level 2, on the other

330

hand, is equally concerned with the transmission of facts, but facts embedded in formal systems of grammar and lexicon requiring much more than mere memorization. News reports or domestic and international events; instructions on how to do or make something; detailed directions for getting to a distant place are excellent examples both of realized Level 2 texts and of the skills required to process them. Thus, well conceived aptitude measures will be designed against the demands of second-language texts of these kinds in any or all skills. The models in current use the Modern Language Aptitude Test (MLAT) developed in 1959 and the DLAB and VORD mentioned earlier work well as predictors of success in learning languages up to Level 2 in the reception mode—reading skills for VORD and most of MLAT and listening skills (among others), for DLAB. (Of the three, MLAT and DLAB were validated decades ago, while VORD, promising in several respects, is still undergoing validation at a government agency). VORD does appear to have an advantage in that persons doing well on this test generally succeed in mastering the syntactic patterns of languages vastly different in structure from English. However, VORD, as noted, is tailored to predicting success in reading only, and in languages employing the Latin alphabet.

The first question to be taken up in the sections below is whether VORD (or any aptitude measure) can forecast attainments beyond Level 2 in any language, and if it can, to which level.

Aptitude Tests as Predictors Beyond Level 2

Existing aptitude measures—singly or severally— have proved to be reasonably satisfactory predictors of success at Level 2 for the three skills in question. But does that suggest a comparable outcome at Level 3 or higher?

There does not seem to be a great deal in the literature of language aptitude testing specifically bearing on success past Level 2, although the desirability of higher levels of attainment is obvious, especially in regard to performance in the work place. The difficulty, though, is to include in the test design those language elements characterizing texts at Level 3 (and higher) as stated or implied in the ILR descriptions. These include references to "...hypothesis, argumentation and supported opinions," the language of which is likely to be relatively rich in lexicon and culturally sensitive. Current aptitude models do not reflect these features; in fact, it is difficult to see how they could be built into tests in the (relatively) short time provided for their administration. It would seem that there are only a few alterna-

136

tives for devising such instruments: a battery including a current aptitude model accompanied by a measure or measures to elicit cultural sensitivity or other psychological aspects; an extended "pure" aptitude test in which, say, Level 3 tasks are embodied in a sophisticated syntax and lexicon, the mastery of which would be exceedingly time-consuming, quite possibly to the point of impracticality.

Prediction of higher level success is but one of the goals of language aptitude testing. Another is the design of a test or test battery which indicates whether the channel of communication (i.e., through the eye or the ear) makes an essential difference. There is no doubt that individuals have preferences in this regard and that memory retention may be a problem for the listening channel, especially if passage replay is not permitted. However, since the item structures in two of the measures are confined to very short language segments at the clause level and below (MLAT and DLAB) the memory load may not seriously affect test performance. VORD on the other hand does include longer texts, especially the ones with the planned blanks (CLOZE-like texts), but, as already observed, it was not designed to test listening comprehension at any point. (For a study of the relationship between MLAT and VORD, see Parry & Child, 1990).

Interestingly, though, VORD at least seems to have some "cross-over" predictive value. Preliminary analyses have been carried out by two government agencies on the performance of language learners on VORD vis-a-vis their subsequent levels of attainment in multi-skill language courses. On the surface, such a result may seem improbable, since speech has a tempo, pattern of pitch or stress, and on occasion tonality, poorly replicated if at all in the writing system (which of course has its own peculiarities often unrepresented in speech). However, language is ultimately a question of communication (very possibly limited to Level 2 for present purposes), so that expectancy based on knowledge of subject matter or familiarity with a situation (sometimes referred to as "semantic feedback") can override problems created by a difference of channel. Thus, VORD, and the other aptitude models as well, may have intrinsic features which allow for cross-channel inferencing. Much more investigation will be required to determine the validity of this hypothesis.

Distances Between and Among Languages

The entire "language aptitude" enterprise could falter in the absence of a comprehensive overview of similarities and differences

1368

among the major languages of the world. There have been over the years a number of attempts to categorize languages in terms of their presumed difficulty; which is to say, how hard they are to learn for native speakers of English. Several of these efforts have in fact been officially blessed within a number of government agencies because they have a certain face validity and have proved useful as general guidelines. However, they do not specify what features of which languages can be expected to cause trouble for learners and which are similar to, or not very different from comparable English features. To lend greater precision to a "global assessment" system it is necessary to determine which major linguistic features of the so-called "hard" languages make learning problematic for English speakers and which lend themselves to (relatively) easy transfer into English.

The following paragraphs set forth what are generally agreed to be the major components of language (whether in speech or writing); the matrix at the table may be useful in following the explanations: (A) phonology (with provision made for written representation); (B) grammatical system, covering what have been traditionally called morphology and syntax; and (C) semantics, taking in meaning in all of its textual representations. These three phenomena are then ordered in such a way as to indicate relative distances of foreign languages from English: Near (1), Middle (2), and Remote (3).

How does all of this relate to language aptitude? The answer is, in a fundamental way, that learning difficulty is tied to the degree in which the object of learning resembles something already known. In the present case it is not enough to say that language X differs greatly from English without specifying the nature of those differences. Let us consider some examples.

A need arises to train a number of individuals in German, a language historically related to English. These persons have had either Spanish or French for one or two high school semesters, not enough to give them sufficient language-learning experience to get off to a fast start in German. That language offers sufficient difficulty to warrant aptitude testing, but is not so daunting overall as to demand very high linguistic skills. Specifically, German uses the Roman alphabet and does so in a way that is roughly isomorphic with the spoken language, hence it is a "near" relative to English (Distance 1). The grammatical system, while sharing many features with English, is sufficiently different syntactically to warrant a "middle"

1363

Table

*Elements of Textual Distance from English to Language to be Acquired**

Textual Elements	Distances		
	Near (1)	Middle (2)	Remote (3)
Writing System as representation of spoken language (A)	•Language in Roman alphabet, with reasonably close fit between "letters" and "sounds." Greek and Cyrillic alphabets may also be included, as they also have close relationships with the associated spoken forms.	•Letters in Roman alphabet at some removal from "sounds" to be repre sented. •Letters in non-Roman alphabet subject to considerable variation in shape, according to positions in words, but generally calibrated to sound system.	•Non-Roman letters change drastically both in terms of positionin words and in clusters they enter into. •Syllabary of considerable size. •Character set, without correspon-dence with sound system.
Grammatical system as frame-work for commu-nication (B)	Surface syntax and vocabulary contain more features similar than dissimilar vis-a-vis English. "True friends" outnumber false ones and there are relatively few features that cause great difficulties for English speakers. Unit lengths of formal structures (e.g., clause, sentence) not greatly different from those of English.	While many grammatical and lexical features resemble those of English there are many points invoving word order, nominalization and the like which English either lacks or uses in more restricted circumstances than is the case in the stances than is the case in the studied language. Lengths of formal units may vary considerably (e.g., much longer, somewhat shorter) than those of English.	Great numbers of grammatical features totally unfamiliar to English speakers: verb/noun relationships, for example, are not like those of English (or other major West European languages). Clause andsentence structure may be marked by affixes or particles rather than by a system of frequent punctua-tion.
Semantic system as cultural outlook (C)	Considerable cultural overlap between language and English; hence, expectancy of learner not basically misplaced.	Cultural milieu of studied language different in many respects from that of English. Still, culture not totally alien.	Culture vastly different from English speaker's; For example, social distinc-tions for caste, gender and the like are formally present in every kind of text.

* Each language may be described by three sets of alphanumerics, e.g., Chinese might be A3/ B2/ C3.

rating (Distance 2); and, in terms of the semantic system, the language expresses a cultural outlook with much in common with that of (American) English speakers, but enough difference to require, again, a Distance 2. Thus, German may be reasonably characterized in the matrix as A1/B2/C2.

Of the available aptitude instruments, which might be the one of choice in this case? Without going into a detailed comparison of the three—for which there is not enough time in this paper—it may be noted that MLAT and DLAB items are confined to word and phrase segments roughly similar to English in length and part-of-speech category, while VORD has sentence- and paragraph-length items expressed by a syntax truly alien to most language learners. Thus, either MLAT or DLAB are preferable to VORD for German.

Turkish may be taken as an example of a much more difficult language system. It does use the Roman alphabet, but with some additional orthographic devices not found in English. Nonetheless, the alphabet enables close coordination between speech and writing, hence rates the digraph A1 (phonology/written representation relatively "near" in terms of learner difficulty).

The grammatical system is a very different matter. Many constructions which would require in English and most European languages verb tenses—past, present, future, and others—employ nouns derived from verbs. These come as a distinct shock to most American students of the language, especially when they are embedded in sentences of 80 to 90 words, a quite usual occurrence in Turkish. Thus, this system can be properly noted B3 (grammar is "remote" from English and a major source of difficulty).

A similar judgment could also be rendered for the semantic system, even though the grounds for the decision are quite different. Turkish has been greatly enriched by its huge number of lexical borrowings from Arabic and Persian and its incorporation of them into the complex Turkish culture. This very complexity can and does cause difficulty to neophytes especially at Levels 2 (high) and 2+ (toward the end of the spectrum in which aptitude testing is likely to be effective). A rating of C, "remote," may be applied here as well. In short, the overall characterization for Turkish is A1/B3/C3.

A final example might be Japanese. This language depends upon both a syllabary (i.e., a consonant plus vowel representation of speech in the writing system) plus a vast number of characters taken over from Chinese. The latter, in addition to imposing a huge memory burden, are not systematically aligned with spoken language segments, with the result that the phonological component, as such, blends

with the lexical (a situation which also pertains to some extent in Korean). The only possible description is "phonology/writing system remote from English," or A3.

Japanese grammar is formidable as well. The verbal system has two basic tenses, but a great number of forms expressing feelings and attitudes of the speaker. These, too, are mostly alien to English, hence, "remote" to the struggling learner. B3 is an apt characterization for the verbal system as well as a number of other Japanese grammar phenomena.

Finally, Japanese culture differs greatly from American culture as expressed in its use of devices reflecting social status. "Polite" vs. "abrupt" verbal forms reflecting particular kinds of social interaction are essential in communication, therefore "musts" for the learner. Again, "remote" is the best description here for semantic differences based on Japanese culture. Thus, Japanese can be captured via the matrix as A3/B3/C3.

The three examples above (German, Turkish, and Japanese) barely scratch the surface of possibilities. They are easily labeled because the characteristic features are so clear-cut. A number of other languages are significantly harder to label: for example, the Cyrillic alphabet used in Russian and several other Slavic languages is problematic for some learners, not so for others. Should it be labeled as A1 or A2?

Such questions are reminders that the designations are, ultimately, relative to a degree and for present purposes reflective of "other languages" considered in the light of English. Literate native speakers of Russian, for instance, would have no problem with the Cyrillic alphabet as used in certain other Slavic languages (though they may have difficulties when it is employed in non-Slavic languages). Nor would a literate native speaker of Turkish have great difficulty with the syntactic patters of Hungarian or Mongolian which in many aspects resemble those of his own language.

References

Carroll, J. B., & Sapon S. M. (1959). *Modern Language Aptitude Test* (MLAT). New York: The Psychological Corporation.

Child, J. R. (1973). *VORD*. Washington, DC: U.S. Department of Defense.

Parry, T. S., & Child, J. R. (1990). Preliminary investigation of the relationship between VORD, MLAT and language proficiency. In T.S. Parry, & C.W. Stansfield (Eds), *Language aptitude reconsidered* (pp 30-66). Englewood Cliffs, NJ:

Prentice-Hall Regents.

Petersen, C. R., & Al-Haik, A. R. (1977). *The development of the Defense Language Aptitude Battery* (DLAB).

Author

JAMES R. CHILD. Senior Research Linguist, National Cryptologic School, Department of Defense. Specializations: language aptitude research, analysis of textual proficiency and performance levels, and translation theory.

1367

Applied Language Learning
1998, Vol. 9, Nos. 1 & 2, pp. 11-30

Zero-Based Language Aptitude Test Design
Where's the Focus for the Test?[1]

Pardee Lowe, Jr.
National Cryptologic School

Aptitude test design has generally led to a "one-test-fits-all" approach, with tests undistinguished as to whether they predict success in language generally without regard to skill modality or whether they are better for one skill over others.[2] Moreover, aptitude test design has failed to distinguish among various levels of possible attainment so that in predicting success one is unable to say whether someone with a high aptitude will generally attain Interagency Language Roundtable (ILR) level 3 in speaking just because she or he has an aptitude for learning a language; nor can one say in what language or type of language. At least one aptitude test—VORD (not an acronym)[3]—attempts to identify those with an ability to learn languages with more complicated grammatical structures like Russian or Japanese. However, its designer, James R. Child, points out that VORD does not predict well the ability to learn tone languages like Chinese. This article looks at the many questions aptitude test design has failed to answer and queries whether it would be possible to return to ground zero (zero-based test design) and design tests that could address them. The questions raised here have not generally been talked about in the literature, but are routinely asked by managers of U.S. Government language programs; so, while sometimes naive, they are actual and important to the success of the U.S. Government's language training effort.

Previous Tests and Their Components

There are several language aptitude tests: The *Artificial*

1368

Language Aptitude Test (ALAT), The *Defense Language Aptitude Battery* (DLAB), The *Elementary Modern Language Aptitude Test* (EMLAT), The *Modern Language Aptitude Test* (MLAT), The *Pimsleur Language Aptitude Battery*, and *VORD*. For government use, three tests, however, should be removed from any serious consideration at the outset: the EMLAT and the *Pimsleur Language Aptitude Battery* that target a lower age group than that of government workers, and ALAT (derived from the MLAT) which is too Eurocentric in design and was replaced at least at the Defense Language Institute Foreign Language Center (DLIFLC) 25 years ago by the DLAB. In other words, the models in current use are instruments better designed for government needs.

While the results from the three remaining tests, DLAB, MLAT, and VORD, have been useful to a degree, the government would like more accurate predictors. The MLAT as administered at the Foreign Service Institute (FSI) of the Department of State correlates consistently around .50 with end of training scores (reported on the ILR scale), and, therefore, accounts for 25% of the variance in student learning.[4] I believe that one needs to characterize aptitude comprehensively and to build more accurate instruments for identifying the ability to learn and to use the language.

We have listed the tests. What then of the tasks they employ to determine language aptitude at least as conceived at the time of test design? An overview of the components of various aptitude tests reveals a divergence of opinion among aptitude test designers on the factors and tasks which contribute most strongly to the construct "language aptitude" (see Table 1).

There seems to be disagreement about what constitutes the construct "language aptitude" itself. We would amplify the phrase "language aptitude" by the words "ability to learn and to use language." While this definition is workable, it by no means replaces a need to more fully define and characterize the ability generally referred to as "language aptitude." The fact that past aptitude test designers have drawn on rather divergent predictors to determine aptitude for learning another language suggests either that there is no agreement on what constitutes the construct "language aptitude," nor that there exist numerous possible predictors that could serve as components in determining the construct, or even that the focus of past test designs has been clear.

One of the most important tasks of the language aptitude testing discipline is to try to reach agreement on a definition of language aptitude. Such a definition should be broad enough to cover the challenge of learning a new language system per se and the ways that system

1359

Table 1

Aptitude as Previously Defined

Sections	Pimsleur	EMLAT	MLAT	DLAB	VORD
Bio Section				X	
Finding Rhymes		X			
Foreign Language Grammar				X	
Grade Point Average	X				
Hidden Words		X			
Interest	X				
Language Analysis	X				
Matching Words		X			
Nouns in the Artificial Language (AL)					X
Number Learning		X	X		
Paired Associates		X	X		
Phonetic Script			X		
Phrases & Sentences in the AL					X
Recognition of Stress Patterns				X	
Sound Discrimination	X				
Sound Symbol Association	X				
Spelling Cues			X		
Text with fill-in-the-blanks in AL					X
Verbs in the AL					X
Vocabulary	X				
Words in sentences			X		

1370

1370

is reflected in natural texts. Reaching broad concurrence on the meaning of aptitude would ultimately result in the production of more accurate tests.

Points for Consideration

Since the times of the writing of the aptitude tests listed above, our concept of aptitude has undergone a radical transformation, more specifically, an expansion. Motivation, learning styles and strategies, as well as teaching styles and methods are part and parcel of the language aptitude picture for many researchers and test designers today. Just how these aspects of language aptitude interrelate is unclear, but Figure 1 suggests a possible view which regards language aptitude as having a core (beyond what it has traditionally contained as the questions later

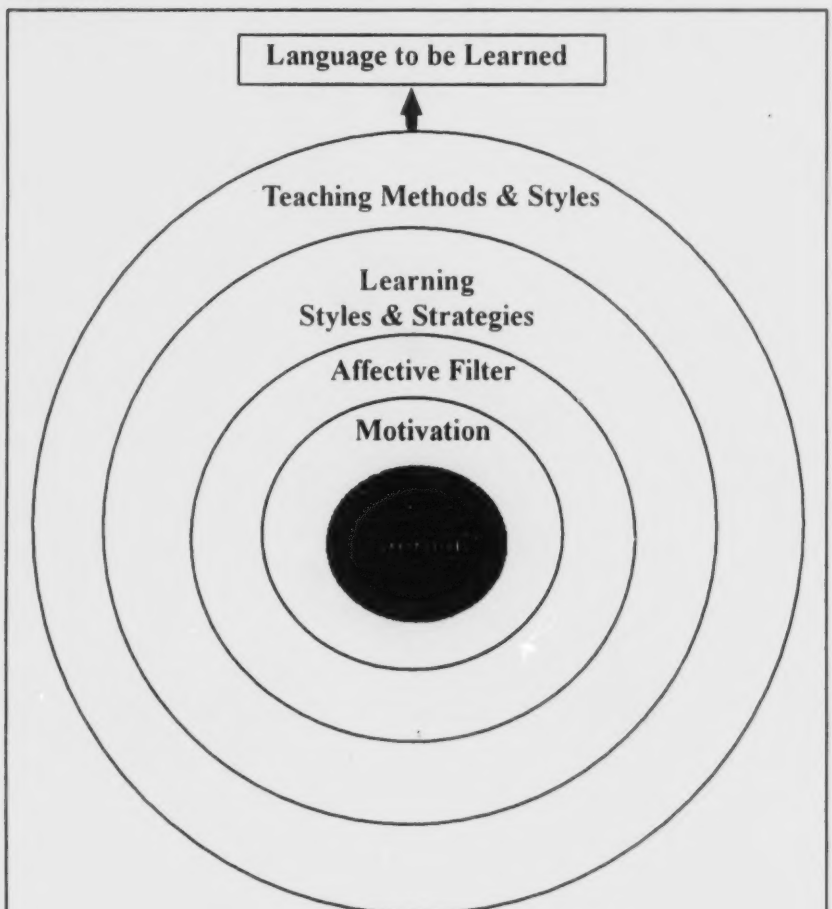


Figure 1. Aptitude and its facilitators

in the article make clear), but which also regards language aptitude as facilitated by other factors, such as motivation, the affective filter and learning styles and strategies on the part of the learner and teaching *methods* and *styles* on the part of the teacher. Their ordering is conceived from the learner's perspective first and the teacher's second. Aptitude resides at the core, but it may be potentiated or attenuated by motivation. We know that high motivation often overcomes low aptitude, while low motivation can defeat high aptitude. On the other hand, a low affective filter allows the full thrust of aptitude and motivation to work, while a high affective filter impedes that combination. Other facilitators depending on the suitability of each to the learner are his learning styles and strategies and the teaching methods and styles of his teacher. Discussion of the interrelationships and contributions of the respective factors leads to an expanded perspective on aptitude.

In this positive ferment, however, there exists a concern that one is failing to focus on some issues pertaining to aptitude test design. This article is concerned with the logical focus of such a design and steps back to regard it from a visual, holistic perspective. Rather than assuming the components employed in aptitude tests so far or the construct that underlies them are adequate, I clear the slate — adopting a zero-based test design. The term *zero-based* is derived from justifying budgets and programs from the ground up; that is, assuming that nothing is sacred and that everything must be re-justified in order to receive funding.

The adoption of the term *zero-based* calls for starting over with new hypotheses. Therefore, I will look at how tests have been designed in the past and ask whether future tests should be designed quite differently, especially in regards to the questions that government managers ask of those who propose to use aptitude tests to identify government employees who could best learn another language. I am not sure that questions government managers pose regarding aptitude test results are always reasonable, nor that one test design can respond to them all. But they form a new point of departure. Moreover, these questions lead to another, perhaps more overarching one:

Q0: What Is to Be the Focus of the Test?

This may be the single most crucial question language aptitude test designers face. Of course, earlier language aptitude tests attempted to define and operationalize each designer's concept of the construct. An approximate .50 correlation between aptitude test scores and exit proficiencies suggests that one might be able to do so more fully.[4]

Moreover, our understanding of the construct has expanded. Finally, the real world poses some rather different questions about language aptitude today than were asked in the past. To answer the question about focus more fully, we list questions an aptitude test could answer in a government context, and we provide background to these questions. These are actual questions government managers have asked when aptitude test results have been used to determine who benefits from training.

A metaphor comes to mind here, that of a runner approaching a course of hurdles. In the future aptitude test we envision here, each question asks if the runner can leap the next hurdle and move on. The test should answer each question as clearly and accurately as possible. These questions doubtless derive from a very broad concept of the construct "language aptitude," one broader than current tests are designed to tap.

Q1: Can an Aptitude Test Tell Us That Someone Can Learn a Foreign Language?[5]

Each year the government devotes significant resources, human and monetary, to training its employees including the military in learning foreign languages. With diminishing resources, the question, "Who is most likely to succeed?" grows more acute.

Q2: Do Languages Have "Personalities," and Is It Possible to Match Language and Person; What is the Effect of Such a Match?

This question, referring to motivation, is not as well addressed in test design as it might be. It asks not just who has a good likelihood of learning a foreign language, but does that person have an interest in that language. Does it match his personality? Does the country, the culture, the ethos of its people, the way they think and act fascinate him? Here the type of motivation plays a role: instrumental, integrative, assimilative? Can a test identify these matches and their strengths?

Q3: How Difficult a Language Can the Examinee Handle?

Early in providing language training, the U.S. Government discovered that Americans learned some languages like Italian and Swedish more readily than they learned German, Russian, or Turkish. In other words, the latter took longer for Americans to master (see Table 2). To deal with this discovery, the government has set aside a longer period for training students to acquire these tongues, allowing

1378

approximately 6 months for the easiest, 9-12 months for somewhat harder ones, and almost 30 months (often including a year in a country) for the hardest. Can one, therefore, predict who will be most successful, not just learning a language, but learning one of a specific category?

Q4: What Language Type(s) Can the Examinee Most Likely Master?

The division of languages into difficulty groups (see Table 2) aids in planning training, but it clusters together languages whose common features may cause Americans difficulties in learning, yet whose nature can differ radically in structure and thought patterns from language to language. Thus, while languages can be grouped together depending on how much time they demand (the current basis of the government's lists grouping languages for difficulty), languages can also be grouped depending on the kinds of difficulties they involve. For example, the four hardest languages for native English-speaking Americans to learn— Arabic, Chinese, Japanese, and Korean— share the difficulty of different writing systems and of non-Western European culture, but from that point on there are more divergencies than commonalities. For example, Chinese is a tone language. Chinese, Japanese, and Korean make a wide variety of sociolinguistic distinctions. Moreover, some languages are agglutinative, like Eskimo or Turkish, in which *evlerden* = *ev* + *ler* (plural marker) _ *den* "from" (Preposition) = "from the houses," others isolating (analytic) like Chinese or Vietnamese, still others inflecting (synthetic) like Slavic languages.[6] Numerous taxonomies could be devised to categorize these differences. The point is that identifying the level of difficulty is not sufficient. Do some learners have an affinity for, say, languages that use noun compounds versus those that use prepositional phrases to describe the same object, e.g., *Schreibmachine* (German), *skrivmaskin* (Swedish), *ritvél* (Icelandic) versus *machine à écrire* (French), and *máquina de escribir* (Spanish), all meaning "typewriter"? What about patterns of thought? With French striving for clarity and concision, while German essays the overarching sentence that perfectly qualifies the topic before it is finally mentioned, we have two truly differing ways of writing about the world.

Q5: In What Skill Modalities Will the Examinee Excel?

In other words, shouldn't aptitude tests be sensitive to modalities? Not everyone writes his or her native language as well as he or she speaks it. Nor does everyone in the world who speaks a language, read it. These facts suggest that while there are four skill modalities—

374

Table 2
DLIFLC Categories of Language Difficulty (Selected Languages)

Category 1	Category 2	Category 3	Category 4
Afrikaans	German	Albanian	Arabic
Danish	Hindi	Amharic	Chinese-
French	Idonesian	Armenian	Mandarin
Haitian-	Malay	Azerbaijani/	Cantonese
Creole	Rumanian/	Azeri	Japanese
Italian	Moldavian	Bashkir	Korean
Norwegian	Urdu	Belarussian	
Portugese		Bengali	
Spanish		Bulgarian	
American		Cambodian	
Caribbean		Czech	
Castilian		Estonian	
Creole		Finnish	
Swahili		Georgian	
Swedish		Greek	
		Hebrew	
		Hungarian	
		Kazakh	
		Laotian	
		Latvian	
		Lithuanian	
		Macedonian	
		Persian-Farsi	
		Polish	
		Russian	
		Serbian-Croatian	
		Slovenian	
		Somalian	
		Tadzhik	
		Tagalog	
		Tatar	
		Thai	
		Turkish	
		Turkmen/Turkoman	
		Ukrainian	
		Uzbek	
		Vietnamese	

37875

Listening, Reading, Speaking and Writing—one person may not control all equally well. Which skill modalities does the government need? How well will a given student learn them? Should a reticent student be trained to speak? A voluble one to write? And so on. Can an aptitude test provide any clues about which skill is most likely the examinee's best or worst?

Q6: How Well Will the Examinee Attain the Course Goals?

Generally, government language courses have a goal and the jobs of graduates have a designated level to show how well the language must be controlled to do the job. For example, at DLIFLC exit scores of 2 in all skills but writing are the minimum target for its basic language programs; at FSI the desired goals are 3 in Speaking and 3 in Reading; while at the Central Intelligence Agency (CIA) a 3 in Listening, Speaking, and Reading are the desired outcomes, with Speaking paramount. Will the student achieve these levels? Can the test provide any predictive clues? Can one predict when a person will plateau? How often will they plateau? Can one predict the frequency of such plateaus?

A further complication in answering these questions stems from a language's relative difficulty for American-born English speakers. The difficulty list masks another problem: namely, language difficulty may change depending on the level we wish the examinee to achieve.

Spanish is often viewed as an easy language. In the government lists it appears as a Category 1 language; that is, among the easier to learn. Yet to put imperatives in Spanish in the negative requires subjunctive forms, and many of the tasks at Level 3 require both the present and past subjunctive. An even more striking example is Indonesian, a Category 2 language. This language is regarded among the easier for Americans to learn. However, few Americans cross the 2+/3 border. Why? A closer look suggests that low on the ILR scale Indonesian is relatively easier to learn, but harder to acquire in the mid and upper ranges. This appears to be due to two interconnected features, morphology and a syntactic feature known as the object-oriented sentence. Neither hurdle is insurmountable, but the main obstacle appears to be that Americans have difficulty knowing those contexts where a native Indonesian speaker would employ the object-oriented sentence.

In other words, besides the morphology and the object-oriented syntactic pattern, the non-native must acquire a new and rather pervasive way of thinking. As a consequence, Indonesian might be regarded as a Category 1 language if one wishes to achieve any level between 0 and 2+, but proves to be a Category 2 language if one wants

to go beyond 2+ into 3, 3+, 4, 4+, or 5! Can an aptitude test predict success for languages of shifting degrees of difficulty such as Indonesian?

Q7: How Far Can a Person Ultimately Go in Learning a Language?

While there may be an immediate exit goal for the course, what indicators can an aptitude test furnish about what the ultimate end point in the person's life-long learning of the target language? In other words, for what level, in what skill, and in what language does the examinee have aptitude?

Most of these questions target areas of ability beyond the information current aptitude tests provide, and, consequently, they extend beyond what has previously been construed as "language aptitude." While these questions have not generally been posed outside government, they logically follow from an expanded construct of language aptitude. If you can learn a language (Q1), what language best matches your personality (Q2), how difficult can it be compared to American English (Q3), and connected with what type of language: isolating, agglutinative, inflecting (Q4), in what skill modalities (Q5), to what level in the course (Q6), and ultimately to what level later in life (Q7)?

An analytic approach might regard these as seven separate questions targeting separate abilities. But a holistic view regards them as ever more precise formulations of the predictive power of a single overall ability — "to learn and to use another language." How can one design tests that will take us further along the path(s) of answering these questions than current language aptitude tests do?

Visualizing Current and Future Aptitude Tests

Visual representations of the focus and the search may be helpful in clarifying what current and future tests may require that is different or that differ from their configuration of tasks. To this end, the following figures posit differing possible solutions to aptitude test design. Although we discuss such designs in a highly abstract manner, we make reference to extant tests.

First, we simplify matters by looking at a single task and its effect on determining aptitude. This is simpler than the reality, but permits us to sketch several approaches in a concise manner before introducing reality's complications.

Let us suppose that ascertaining a person's ability to learn and use a foreign language were determinable by a single task and that task would target the speaking skill modality. We choose speaking because it is learned in the native language before reading and writing, and moreover, because it is a production skill (see Figure 2). How would we proceed? Could a speaking test in one's language determine ability to learn and use a second language? If so, what skill modality?

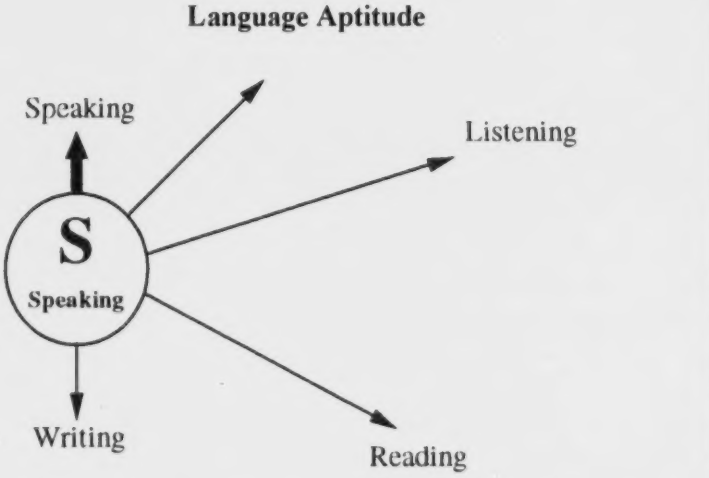


Figure 2. "One size fits all": One skill modality is the predictor

While we are operating here with single tasks, the points made about them could be readily expanded to include any range of tasks for the skill modality the single task represents. Hence, instead of a single speaking task, we could conceive of a battery of two or more tasks, and extrapolate results from the battery rather than from the single task.

Assume for a moment, however, that we found the one speaking task telling us that a person could speak well. Note it would indicate some ability to learn a language. But would it predict the ability to read, to write? Presumably it would include the ability to listen. But how well? (See Lowe, 1985, for the "offset" problem.) Since people learn to read and write after they learn how to speak and to listen, and some never learn to read, and still others never learn to write, using a single speaking task won't work! Obviously a single task could not predict in those cases where the other skill modalities, reading and writing, were never learned! Moreover, even if the chosen task could predict an

ability to learn and to use those skill modalities, the "offset" problem though not fully worked out for all the skill modalities would resurface. Think of the "offset" between one's ability to speak, listen, and read, versus one's ability to write. The last, in the U.S. at least, is usually lower than that of the other skills. So a single speaking task might indicate an ability to learn to speak and indicate some "general language ability," but it might well not predict ability to learn the other skills.

What about a single "general language factor," Oller's unitary competence hypothesis? Oller based this hypothesis on the "g" factor in intelligence. Oller has been inclined to equate the two, "g" and a "general language factor (glf)." Subsequently, he withdrew the hypothesis, having learned that his proof failed for various statistical reasons (Oller, 1980, cited in Vollmer, 1983). Still, I separate them in this article.

The unitary competence factor, however, has continued its existence anyway (Hughes & Porter, 1983). Why won't it die? Well, it is possible to adduce evidence for Oller's position up to a point, indicating that there is doubtless a strong glf, but that it is simply not as strong as "g" in intelligence. Contemplate the following situation: You give a well-designed listening comprehension test at ILR Level 3. One particular examinee does exceedingly well on the test. Your reaction might well be, not only that he understood the language at the level in question, but that the test in fact indicates that he really knows the language rather well. Yet, at the same time, you could in no way state that he also *writes* the language equally well; that is, tested listening ability cannot predict ability in another skill. You must test separately for that. I have not found it possible to predict language learning ability in a particular skill modality unless the aptitude test contained *tasks targeted against that modality*. Hence, one cannot predict a student's level of reading ability from the level of speaking ability. The import for aptitude test design is that we should have a separate listening, reading, speaking, or writing component or even separate tests, if we desire to make statements about a given skill modality.

There is in our abstract approach a possible single task solution involving "g" or "glf" (see Figure 3). This might well be a cognitive task that would determine general language aptitude and, ideally, ability in the other skill modalities as well. If one equates "glf" with "g," then the task could be drawn from intelligence testing. If one keeps "glf" separate from "g," then the task to be chosen would be taken from those thought to represent "glf" that do not connect with general intelligence tasks, and would therefore predict the examinee's degree of "glf."

1379

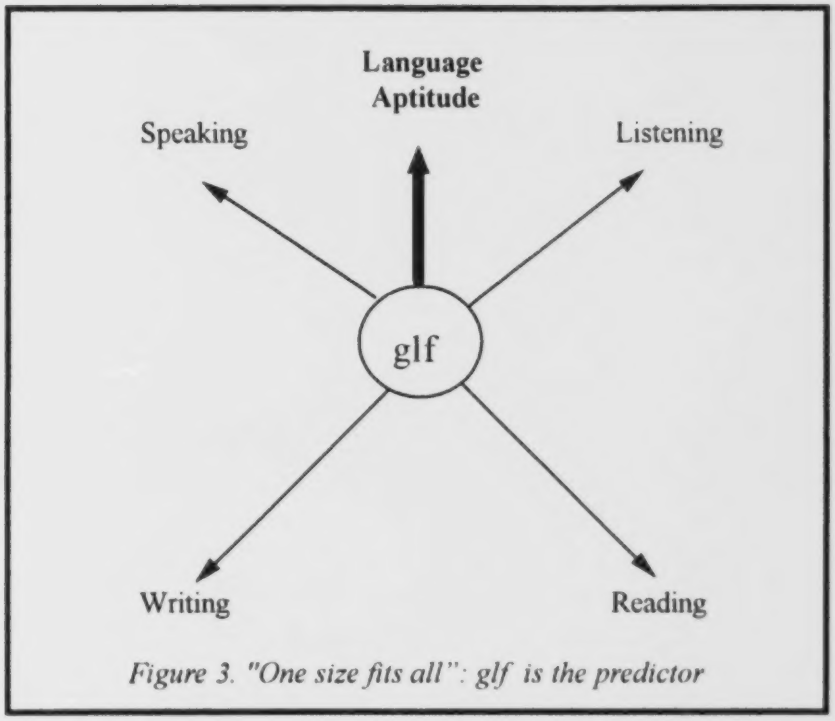


Figure 3. "One size fits all": glf is the predictor

In the total test, tasks representing "glf" need to be supplemented by tasks representing any single skill modality to ascertain the extent of possible success in that skill modality. Every skill modality has a set of related subskills that must be tested in order to ascertain its full ability. Which subskill(s) must be tested is an open question. With these thoughts we leave the "one size fits all" philosophy of ascertaining language aptitude (either that of a single skill modality or that of "glf").

The opposite pole is the "Mixed Menu" philosophy: "One from column A, one from Column B." Staying again with a single task per skill modality, we could have one task for predicting ability, one task for listening, one for reading, one for speaking, and one for writing. (See Figure 4). The question to be answered is: Does the whole equal the sum of the parts? Does predicting success in each of the four skills provide some entrée into overall language aptitude? A variation on this solution would introduce into the previous set of tasks one task for predicting the "glf." Again, does the whole (a fuller whole this time) equal the sum of more of its parts? (See Figures 5 & 6.)

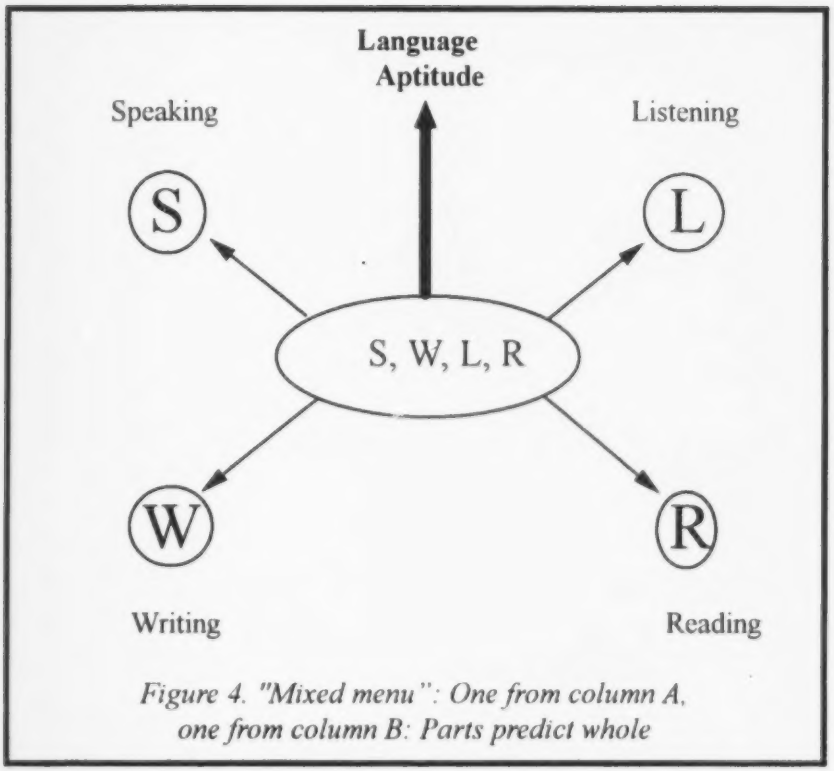


Figure 4. "Mixed menu": One from column A, one from column B: Parts predict whole

1381

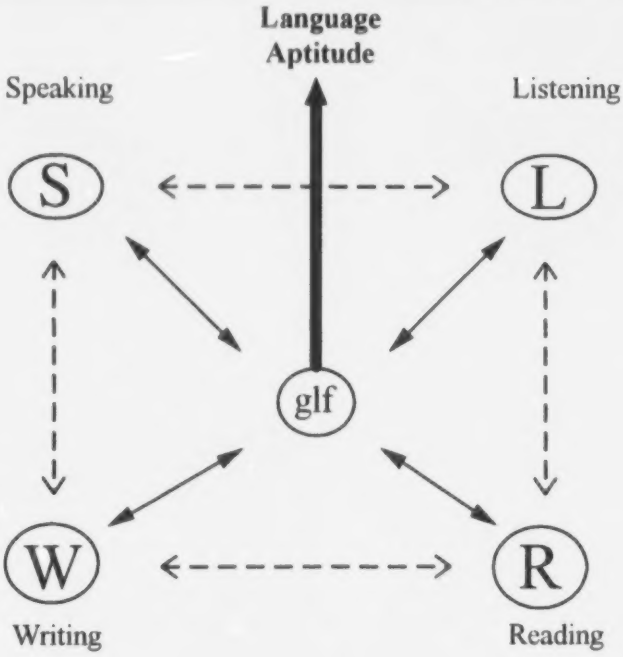


Figure 5. "Mixed menu": One from column A, one from column B

Considering the extant aptitude tests, we can say that their designs more fully resemble the "Mixed Menu" solution, often with multiple tasks. Normally, we choose these tasks by such statistical methods as factor analysis. We try a number of tasks possibly indicating "language aptitude." Then, through factor analysis we identify those that load most on the trait. Subsequently, the ones with the highest loadings are selected for the aptitude test.

In reviewing the current tests, one is led to the questions: What are these tests focusing on? Which of the questions posed earlier does a given aptitude test attempt to answer? With what success?

1352

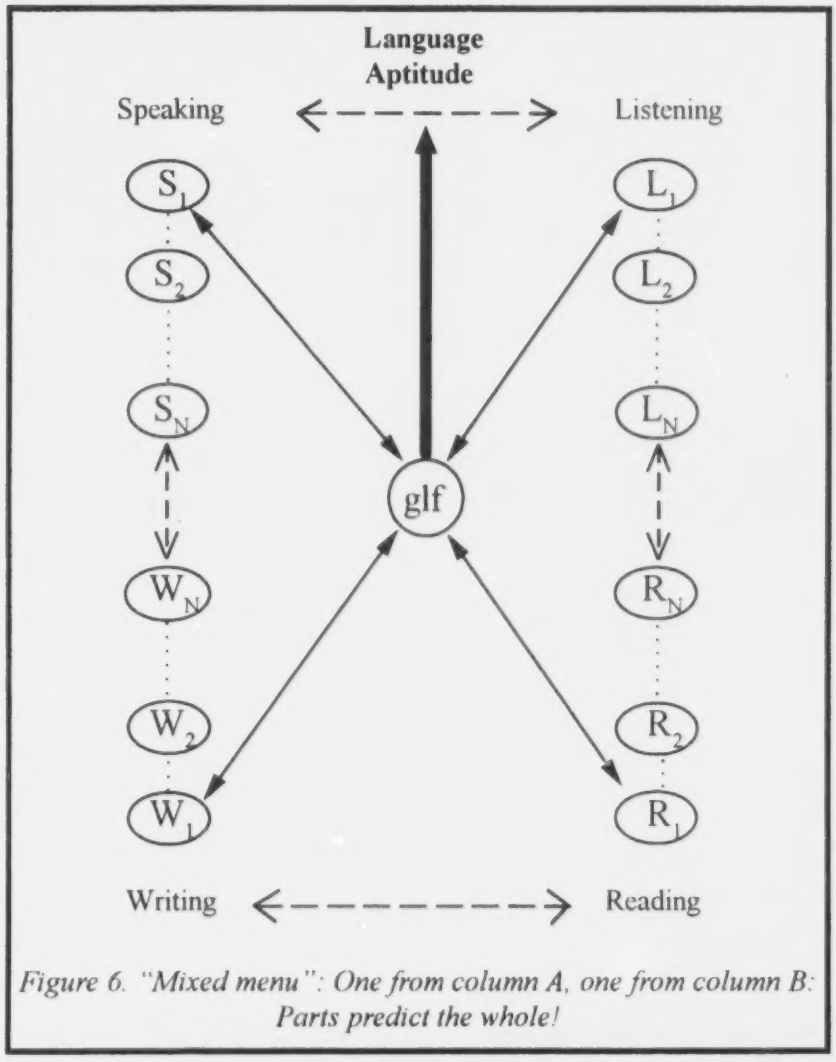


Figure 6. "Mixed menu": One from column A, one from column B: Parts predict the whole!

Current tests appear to answer our first question with some degree of accuracy: i.e., a correlation of .50 between score on the MLAT and exit proficiencies. Table 3, however, suggests that correlations between DLAB scores and exit proficiencies vary for individual skill modalities and test design.This variance again raises the question, what is the focus of the test?

1383

Table 3
Correlations Between DLAB scores and Outcome Variables by Language by Year: Russian

YEAR	STATISTIC	DLPT-L	DLPT-R	DLPT-S
1986	r	0.29077	0.40175	0.12764
	p	0.0001	0.0001	0.0006
	n	729	729	729
1987	r	0.34892	0.41579	0.22475
	p	0.0001	0.0001	0.0001
	n	567	566	567
1988	r	0.29463	0.37770	0.17589
	p	0.0001	0.0001	0.0001
	n	636	636	636
1989	r	0.29064	0.31847	0.22066
	p	0.0001	0.0001	0.0001
	n	834	834	834
1990	r	0.31858	0.32597	0.20292
	p	0.0001	0.0001	0.0001
	n	833	833	832
1991	r	0.31147	0.36340	0.18500
	p	0.0001	0.0001	0.0001
	n	796	796	796
1992	r	0.20546	0.22894	0.17374
	p	0.0001	0.0001	0.0001
	n	824	824	824
1993	r	0.22952	0.28561	0.22778
	p	0.0001	0.0001	0.0001
	n	678	678	678
1994	r	0.28512	0.46500	0.22988
	p	0.0001	0.0001	0.0002
	n	255	255	255

Since the current aptitude tests are paper-and-pencil only tests, they automatically involve reading. Some have audiotaped sections, but do such sections really test listening or is their primary goal to test memory span? All these tests are machine-scorable so they cannot test writing. And none contain any spoken language tasks, so they cannot test the ability to learn speaking! What are we testing? What is the concept of aptitude that underlies the test? What can we say as a result

1384

of the test? Perhaps, in retrospect, we should wonder that we can say as much as we do at least as regards general aptitude.

Conclusion

Focusing on a future aptitude test design, should we:

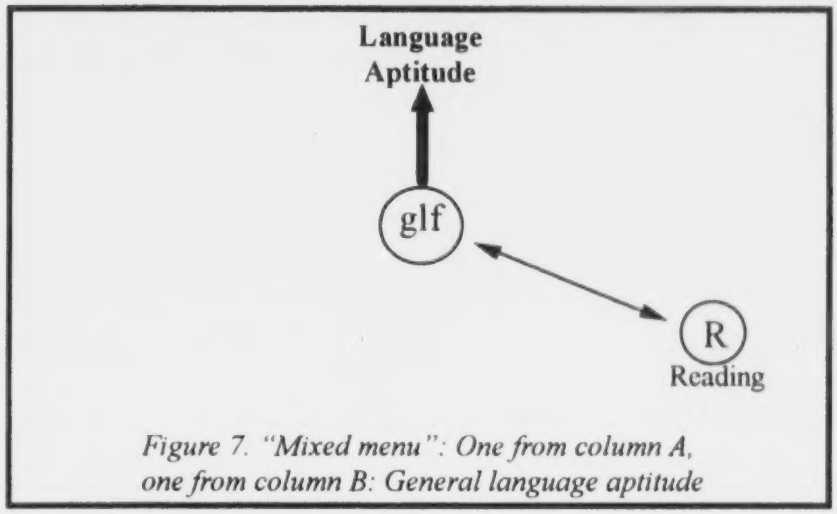


Figure 7. "Mixed menu": One from column A, one from column B: General language aptitude

- attempt to test for "glf" to determine those who would and
- those who would not benefit from language training?
- target aptitude tests against against a single skill modality, for example, reading (see Figure 7)?
- give targeted aptitude tests against languages of a specific structure such as Child's VORD with its Turkic-based artificial language?
- attempt to predict how far the examinee will go in learning the language?

These questions could be multiplied, but their implication is clear. In the future, what will be our definition of the "language aptitude" construct? And more importantly, what will be the focus of our language aptitude tests (Q0)? Which of these questions can a language aptitude test conceivably answer satisfactorily? Which only partially, and which not at all?

3885

Notes

[1] The views expressed herein are those of the author and in no way represent those of the Department of Defense.

[2] I am indebted to both Ray T. Clifford and James R. Child for critiquing earlier drafts of this article. The latter was particularly helpful in the revision of the earlier oral version, from September 1994, into this written one. I regret only that I have been unable to reflect all of their comments. Naturally, any remaining errors are the responsibility of the author. Finally, I also wish to thank John Lett for providing the correlation statistics from ongoing research at DLIFLC.

[3] VORD is not an acronym. Child has been so plagued by questions as to its meaning that he ultimately resorts to saying, "No, it's not an acronym! But if you must confer meaning on it, then I'll say that it's the word for 'word' in VORD."

[4] Ray T. Clifford (through personal communication) points out that "... .50 is actually very high compared to other attempts at predicting human behavior over an extended period of time." My point is rather: Can we say more about less, which I hope will become clear through the designs suggested later in the article?

[5] After Q1 I am not exactly sure as to the order of the following questions.

[6] Bernard Comrie. (1981). *Language Universals and Linguistic Typology.* Chicago, IL: University of Chicago Press

References

Artificial Language Aptitude Test (ALAT). (1951). Washington, D.C.: Department of the Army.

Carroll, J. B., & Sapon, S. M. (n.d.) *Elementary Modern Language Aptitude Test* (MLAT).

Carroll, J. B., & Sapon. S. M. (1959). *The Modern Language Aptitude TEST* (MLAT). New York: The Psychological Corporation.

Child, J.R. (1973). VORD. Washington, DC: Department of Defense.

Comrie, B. (1981). *Language universals and linguistic typology.* Chicago, IL: University of Chicago Press.

Defense Language Aptitude Battery (DLAB). (1977). Monterey, CA: Defense Language Institute Foreign Language Center.

Lowe, Jr., P. (1985). The ILR Scale as a synthesizing research principle. In C. James (Ed.), *Foreign language proficiency in the*

7886

classroom and beyond. Lincolnwood, IL: National Textbook. P. 35ff.

Parry, T. S., & Child, J. (1990.) Preliminary investigation of the relationship between VORD, MLAT, and language proficiency, In T. S. Parry & C. W. Stansfield (Eds.). *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice Hall Regents. Pp. 30-66.

Petersen, C. R., & Al-Haik, A. R. (1976). The development of the Defense Language Aptitude Battery (DLAB). *Educational and Psychological Measurement, 6*, 369-380.

Pimsleur, P. (1966). *The Pimsleur Language Aptitude Battery*. New York: Harcourt Brace Jovanovich.

Vollmer, H. J. The structure of foreign language competence," In A. Hughes & D. Porter (Eds.), *Current developments in language testing*. London: Academic Press.

Author

PARDEE LOWE, JR. Senior Language Analyst, National Cryptologic School, Department of Defense. Specializations: aptitude, oral proficiency testing, passage grading, performance, and proficiency testing.

1387

Applied Language Learning
1998, Vol. 9, Nos. 1 & 2, pp. 31-70

The Modern Language Aptitude Test for Predicting Learning Success and Advising Students

Madeline Ehrman
Foreign Service Institute

The Modern Language Aptitude Test (MLAT) was part of a project examining biographical, motivational, attitudinal, personality, and cognitive aptitude variables among 1,000 adult students preparing for overseas assignments at the Foreign Service Institute (FSI) with various smaller numbers for sub-samples completing different instruments. Data were analyzed by correlation, Analysis of Variance (ANOVA), chi-square, and multiple regression as appropriate to the data and the research questions. The MLAT proved the best of the available predictors of language learning success. As part of an effort to expand the concept of language learning aptitude beyond strictly cognitive factors, this study relates the MLAT not only to end-of-training proficiency outcomes, but also to personality dispositions, using both overall correlational data and information on extremely strong and weak learners. Qualitative findings from use of the MLAT part scores in student counseling activities are also described, suggesting utility for this instrument beyond prediction of learning success.

This article describes findings of research in progress at the Foreign Service Institute (FSI), a U.S. government language training institution. For years, incoming students have taken the MLAT; indeed, a sample from FSI was among the groups on which the MLAT was originally normed (Carroll & Sapon, 1959). It is still used as part of the institute's procedures for assignment to foreign language training. (Language aptitude testing is also done at other agencies.)

© *1998, Madeline Ehrman*

1388

Over recent years, the MLAT has become the subject of some controversy at FSI. Some program managers continue to see a good relationship between performance on the MLAT and in language training. Others protest that the relation, such as it is, is not very strong and furthermore the MLAT may not represent the true ability of those who lack formal education (Rockmaker, personal communication, 1993). Anti-MLAT opinion has also suggested that the MLAT was designed for the audio-lingual methodology that was in vogue in the late 1950s and 1960s and that the test is no longer valid for the much more "communicative" teaching that is now done at FSI (Bruhn, personal communication, 1992). Much of the distrust of the MLAT is connected with the increased suspicion of psychological testing during the last quarter century (Anastasi, 1988). The project on which this paper reports was initiated in order to take such concerns about the MLAT out of the realm of allegation and find out just how useful it still is.

The present article reports on two efforts to addressthese concerns. One is a quantitative investigation of a large sample of FSI students between 1992 and 1994. That study looks at the MLAT primarily as a predictor of language learning success in the FSI setting of intensive, full-time language learning for communicative use. The other portion of the article describes a less rigorous attempt to make use of patterns of high and low MLAT part scores with individual students. The initial outcomes of this attempt, still highly exploratory, suggest that the MLAT may have value for pinpointing areas of learning success and difficulty for a wide range of students, including some relatively able but context-dependent ones not well served by relatively grammar-oriented instruction.

Review of Literature

The MLAT was perhaps the culmination of a long tradition of psychometric test development and efforts to predict language learning achievement. It achieved a fairly respectable level of success in the audio-lingual and grammar-translation classrooms of the 1950s and 1960s (Spolsky, 1995). Other important language aptitude tests developed out of the same tradition include the Pimsleur Language Aptitude Battery (PLAB) (Pimsleur, 1966), the Defense Language Aptitude Battery (DLAB) (Petersen & Al-Haik, 1976), and VORD (Parry & Child, 1990). The Pimsleur is different from the MLAT in particular because it includes a portion directly addressing the ability to infer language structure from an artificial language stimulus. The DLAB

1989

consists primarily of such induction-testing items, in a modified English. VORD was designed to test the ability to cope with the grammar of languages in the Altaic family and consists of items that test such grammatical prowess (Parry & Child, 1990). All four, including the MLAT, were found to have similar predictive validity (Parry & Child. 1990). This article will not address these other instruments, but will focus on the MLAT, which is the instrument that is still in use at the Department of State.[1]

The outcome of a major research project at Harvard University, the MLAT is based on a factor analysis of a large number of individual characteristics thought to contribute to language learning. Carroll (1962) describes the project in extensive detail; the MLAT Manual (Carroll & Sapon, 1959) provides information on the validation studies. The individual characteristics were grouped into four main categories: phonetic coding ability (distinguishing sounds and reflecting them graphically), grammatical sensitivity (recognizing and using syntactic relationships), memory (rote and contextualized), and inductive language learning. All but the last of these four are directly addressed in the five parts of the MLAT (see Appendix A).

Other components listed by scholars of language aptitude include motivation and knowledge of vocabulary in the native language (Pimsleur, 1968), the ability to hear under conditions of interference (Carroll, 1990), the ability to "handle decontextualized language" (Skehan, 1991), and the ability to shift mental set and cope with the unfamiliar (Ehrman, 1994b, 1995b, 1996; Ehrman & Oxford, 1995).

A desire for better prediction of language learning and the ability to exploit aptitude testing further has led to recent research efforts. At least two major projects in recent years have examined the role of individual differences in addition to strictly cognitive aptitude in language learning. They are the Defense Language Institute Foreign Language Center's (DLIFLC) Skill Change Project (Lett & O'Mara, 1990) and FSI's Language Learning Profiles Project (Ehrman, 1993, 1994, 1995b, 1996; Ehrman & Oxford, 1995; Oxford & Ehrman, 1995). Both investigated such variables as biographic factors, personality, motivation, anxiety, and learning strategies, as well general intelligence (DLIFLC only). A similar project was begun at the Central Intelligence Agency language school, though without personality variables, and DLIFLC engaged in a large-scale effort to review the DLAB (Thain, 1992; Lett & Thain, 1994). This article is part of the project at FSI.[2]

Across a number of studies, predictive validity correlations for the MLAT have generally ranged between .42 and .62 for most

1390

languages, with outliers of .27 for certain non-Indo-European languages at the DLIFLC and as high as .73 with language instructor ratings of student performanced at FSI (Carroll & Sapon, 1959). More recent studies of the MLAT produce quite mixed results. Brecht, Davidson, and Ginsburg (1993) did not find the MLAT predictive of overall oral proficiency in intensive language training in Russian. However, for the same programs they found Part 3 (Spelling Clues) to be "highly significant" in predicting listening comprehension and the Total Score to be significantly predictive of reading proficiency. They speculate that the complex nature of the communicative task causes the lack of predictive value for oral proficiency. This suggestion is quite consistent with the questions raised at FSI (see above) and the point of view that standard aptitude measures do not "take into account" such developments as focus on communicative competence, pragmatics and discourse, new thinking by cognitive psychologists (Parry & Stansfield, 1990).

Another finding is that of Spolsky (1995), who reports that MLAT Part 1 correlated significantly with success on the part of Israeli learners of French as a foreign language, but the MLAT did not predict achievement in Hebrew at the same school. He suggests that this variance may be related to differences in such factors as motivation, which is so powerful that it may override aptitude. (I suggest that it may also be the case that the students were learning Hebrew as a second language, not a foreign language, so not all of their learning was classroom-based, which is the task for which existing language aptitude tests were designed.)

Most of the research cited addresses the use of the MLAT (and other aptitude measures) as predictors of learning success, and indeed this is an important consideration for assignment to intensive and long-term language training at taxpayer expense. However, a measure like the MLAT also has potential utility for *placement* in a program (Wesche, 1981) and *diagnosis* of learning difficulties, for counseling students, and for tailoring programs to their needs (e.g., Demuth & Smith, 1987; Sparks, Ganschow, & Patton, 1995). These applications have received far less attention in the literature. They are also among the areas of interest for the FSI investigation, and it is in these that the MLAT has been successfully used (Lefrancois & Sibiga, 1986; Wesche, 1981).

1391

Methods

Sample

In this study, there are 343 students altogether with at least a single MLAT score; of these, part scores for the five subscales are available for 296. Males constitute 59% and females 41% of the sample. The average age of students at the time of participation was 39, with a standard deviation of 9 years. The median education level was between bachelors and masters degrees. Of those that report previous language study, the average number of languages studied was 1.8.

In the presentation of correlations with other instruments, numbers are smaller because not every person in the data set with an MLAT score completed all the other instruments. For example, of the 343 students with at least one MLAT score, only 93 had scores on the Myers-Briggs Type Indicater (MBTI) Form G. On Tables 4 and 5 in the results section of this article, which are excerpted from another sub-study in the FSI Language Learning Profiles Project , the numbers are different from those in the present study, though they represent overlapping sub-samples from the same population of students.

FSI trains and tests students not only from its parent agency, the Department of State, but also from many other agencies. Students from the Department of State comprise 70% of FSI language students. Other agencies sending the most students are the United States Information Agency, the Department of Defense, the Department of Commerce, and the Agency for International Development.

Students in this study are beginners in long-term (i.e., 16 weeks or longer) intensive language training. The languages they are studying are classified into four categories based on agency experience with the length of time needed by English speakers to reach "professional" proficiency (S-3, R-3—see 'Instrumentation' for a brief description of the ILR rating scale):

1. Western European;
2. Non-Western European but relatively quick for English speakers to learn (Swahili, Indonesian, and some North European languages);
3. Other non-Western European (e.g., Russian, Thai), but excluding the Category 4 languages;
4. "Superhard" languages (Arabic, Chinese, Japanese, and Korean).[3]

1392

Usual training lengths vary by language category. Most FSI students are expected to reach "professional" proficiency (S-3 R-3) in 24 weeks in a Category 1 language, in 32 weeks in a Category 2 language, in 44 weeks in a Category 3 language, and in 88 weeks (2 academic years) in a Category 4 language.[4] These expectations are normally reflected in the lengths of student assignments to training and are also accounted for in the statistics reported in this article.

Instrumentation

The MLAT (Carroll & Sapon, 1959) is the classic language aptitude test, with 146 items. The MLAT Manual describes its five parts: 1. Number Learning (memory, auditory alertness); 2. Phonetic Script (association of sounds and symbols); 3. Spelling Clues (English vocabulary, association of sounds and symbols); 4. Words in Sentences (grammatical structure in English); and 5. Paired Associates (memorizing words), together with a total score. The MLAT was correlated .67 with the Primary Mental Abilities Test (Wesche, Edwards, & Wells, 1982), suggesting a strong general intelligence factor operating in the MLAT. Split-half reliabilities for the MLAT are .92 - .97, depending on the grade or age. For college students, validity coefficients (correlations with course grades) provided in the MLAT Manual (Carroll & Sapon, 1959) are .18 - .69 for the long form of the MLAT and .21 - .68 for the short form. For adult students in intensive language programs, validity coefficients (correlations with teacher ratings) in the Manual (Carroll & Sapon, 1959) are .27 - .73 for the long form and .26 - .69 for the short form. This study used the long form.

The subscales of the MLAT are described briefly in Appendix A. The Index Score used at FSI originated in the 1960s as a T-score based on the Total score, with three standard deviations of 10 on either side of a mean of 50.[5] It has since become frozen as a translation of the Total, much like Scholastic Aptitude Test scores until recently, because of the agency personnel system's dependence on over 30 years of Index records. For users of the MLAT who are more familiar with the raw Total score, a table of equivalences is provided in Appendix B.

Note that an Index of 50 is the mean established when the MLAT was originally normed and includes a variety of subjects from high schools and colleges. Whether it in fact is still representative of the population outside FSI is uncertain. What is certain, however, is that a mean Index of 50 is no longer valid for FSI students. There has

1393

been a gradual upward tendency in the MLAT Index mean at FSI over the intervening 30 years. Wilds (1965) reported a mean Index of 54 (N=957, no SD); an agency-internal document reports a 1984 mean Index of 59, SD 10, N-312 (Adams, 1984); and the mean Index for all the students in the current sample who had MLAT scores is 63, SD 10, N = 343. [6]

End-of-Training Proficiency Tests

These tests provide the main criterion measure in this study. At the end of training, FSI students are given proficiency assessments resulting in ratings ranging from 0 to 5 for speaking (the S-score, which includes interactive listening comprehension) and for reading (the R-score). The full oral interview, including speaking, interactive listening, and an interactive reading test using authentic material, takes two hours. R-3, for example, indicates reading proficiency level 3 ("professional" proficiency); S-2 represents speaking proficiency level 2 (working proficiency). Other levels are 0 (no proficiency), 1 (survival level), 4 (full professional proficiency, with few if any limitations on the person's ability to function in the language and culture), and 5 (equivalent to an educated native speaker). "Plus" scores (e.g., indicating proficiency between S-2 and S-3) were coded as 0.5; thus, for example, a score of S-2+ was coded 2.5.

The ratings are equivalent to the guidelines of the ILR/ACTFL (Interagency Language Roundtable/American Council on the Teaching of Foreign Languages) that originated at FSI and have been developed over the years by government agencies. These guidelines are detailed by Omaggio, 1986. Most students enter FSI with goals of end-of-training proficiency ratings at S-3 R-3 for full-time training, comparable to ILR/ACTFL Superior Proficiency.

Reliability studies have shown that government agencies have high inter-rater reliability for proficiency ratings within a given agency, but that the standards, or their interpretations, are not always the same at every agency. Thus, raters at different government agencies do not have as high an inter-rater reliability as raters at the same agency. Proficiency ratings are therefore considered reliable indicators of the level of language performance of an individual student within an agency (Clark, 1986).

1394

Learning Style, Strategy, and Personality Instruments

The Learning Style Profile is a pure learning style instrument: that is, it is neither a personality questionnaire nor an aptitude test. The Myers-Briggs Type Indicator and its Type Differentiation Indicator scoring system are both a personality instrument and a way to assess learning style, as is the Hartmann Boundary Questionnaire. The student learning activities questionnaires tap learning strategies.

The Hartmann Boundary Questionnaire (HBQ, Hartmann, 1991) was developed for research with sleep disorders and nightmares, using a psychoanalytic theoretical base. It is intended to examine the degree to which individuals separate aspects of their mental, interpersonal, and external experience through "thick" or "thin" psychological boundaries. Its 146 items address the following dimensions: sleep/dreams/wakefulness, unusual experiences, boundaries among thoughts/feelings/moods, impressions of childhood/adolescence/adulthood, interpersonal distance/openness/closeness, physical and emotional sensitivity, preference for neatness, preference for clear lines in pictures or clothing, opinions about children/adolescents/adults, opinions about lines of authority, opinions about boundaries among groups/peoples/nations, opinions about abstract concepts, plus a total score for all twelve of the above scales. For example, thin boundaries are represented by commonly drifting in and out of sleep states while waking up, memory for experiences at a variety of ages, tolerance for lack of order in the workplace, or preference for little organizational hierarchy. Thick boundaries are suggested by the opposite approach, e.g., preference for thick, heavy clothing, interpersonal distance, or beliefs that children should be seen and not heard. Hartmann found women and younger people score consistently "thinner" than men and older people. Cronbach alpha reliability for the HBQ is .93, and theta reliabilities for subscales are .57 - .92 (Hartmann, 1991).

The National Association of Secondary Schools Principals' Learning Style Profile (LSP), (Keefe & Monk, with Letteri, Languis, & Dunn, 1989) is a 125-item composite measure composed of many different approaches to measuring learning style. The main subscales are cognitive skills (analytic, spatial, categorization, sequential processing, detail memory, discrimination), perceptual response (i.e., sensory preferences: visual, auditory, emotive/kinesthetic), orientations (persistence, verbal risk-taking, manipulative), study time preferences (early morning, late morning, afternoon, evening), and environmental context for learning (verbal vs. spatial, posture, light, temperature, mobility, and

1393

grouping). Cronbach's alpha for the subscales ranged from .47 to .76, with an average of .61. Test-retest reliabilities were .36 to .82 after 10 days and somewhat lower after 30 days. Concurrent validity of the LSP's analytic subscale with the Group Embedded Figures Test was .39. Concurrent validity of the perceptual response subscales of the LSP with the Edmonds Learning Style Identification Exercise was .51 - .64. Many of the environmental context subscales of the LSP correlated with Dunn and Dunn's Learning Style Inventory, .23 - .71. All concurrent validity scores are reported in the manual with a significance value < .002.

The Myers-Briggs Type Indicator (MBTI, Myers & McCaulley, 1985), Form G, is a 126-item, forced-choice, normative, self-report questionnaire designed to reveal basic personality preferences on four scales: extraversion-introversion (whether the person obtains energy externally or internally), sensing-intuition (whether the person is concrete/sequential or abstract/random); thinking-feeling (whether the person makes decisions based on objective logic or subjective values); and judging-perceiving (whether the person needs rapid closure or prefers a flexible life). Internal consistency split-half reliabilities average .87, and test-retest reliabilities are .70 - .85 (Myers & McCaulley, 1985). Concurrent validity is documented with personality, vocational preference, educational style, and management style (.40 - .77). Construct validity is supported by many studies of occupational preferences and creativity.

The Type Differentiation Indicator (TDI) (Saunders, 1989) is a scoring system for a longer and more intricate 290-item form (MBTI, Form J) that provides data on the following subscales for each of the four MBTI dimensions: extraversion-introversion (gregarious-intimate, enthusiastic-quiet, initiator-receptor, expressive-contained, auditory-visual); sensing-intuition (concrete-abstract, realistic-imaginative, pragmatic-intellectual, experiential-theoretical, traditional-original); thinking-feeling (critical-accepting, tough-tender, questioning-accommodating, reasonable-compassionate, logical-affective); and judging-perceiving (stress avoider-polyactive, systematic-casual, scheduled-spontaneous, planful-open-ended, methodical-emergent). The TDI includes seven additional scales indicating a sense of overall comfort and confidence versus discomfort and anxiety (guarded-optimistic, defiant-compliant, carefree-worried, decisive-ambivalent, intrepid-inhibited, leader-follower, proactive-distractible), plus a composite of these called "strain." Each of these comfort-discomfort subscales also loads on one of the four type dimensions, e.g., proactive-distractible is also a judging-

1396

perceiving subscale. There are also scales for type-scale consistency and comfort-scale consistency. Reliability of 23 of the 27 TDI subscales is greater than .50, an acceptable result given the brevity of the subscales (Saunders, 1989).

Student Learning Activities Questionnaires

At the beginning of training, students were asked to complete the Motivation and Strategies Questionnaire (Ehrman & Christensen, 1994), which has a variety of questions about self-efficacy as a learner, motivation, anxiety, a detailed list of classroom activities, and a set of individual activities. At the end of the students' training (minimum of 16 weeks, maximum of 44 weeks, depending on the difficulty category of the language), each participant in the study was asked to complete two questionnaires: "CLASSACT" (Ehrman & Jackson, 1992) on relative usefulness of a fairly detailed list of classroom activities (Likert scaled 1-3) and "SELFACT" (Hart-Gonzalez & Ehrman, 1992) on relative usefulness (1-3) of their own study activities and estimated time per week devoted to each. These questionnaires are used here for the first time. Because completion at the end of training was voluntary and students were very busy with preparations for departure, the return rate was low (about 10%), and numbers for several of the items are not adequate for analysis. (This and other studies using these questionnaires are part of their validation. When there are sufficient cases, they will be subjected to reliability analysis and factor analysis.)

Data Collection and Analysis

Data collection took place over a two-year period, between 1992 and 1994. Students who started at the beginning of each semester were asked to participate but could decline the invitation; under 5 % of the students who were approached chose not to participate. During the 1992-1993 academic year, all French and Spanish students (who start 10 times a year) were also invited to join the study, with the same declination rate.

All questionnaires except the MLAT were administered within the first week of training. If a student already had an MLAT record, he or she could arrange for those scores to be included in the research data set; otherwise, MLAT administration took place within the first month of the beginning of training. In this sample, almost all (95%) of the MLAT scores were current, i.e., within the previous 3 years. Proficiency tests were administered at the end of training, after (in most

1397

cases) 24 or 44 weeks.

Data analysis in this study on Statistical Product and Service Solutions (SPSS) for Windows 5.0.1 (Norusis, 1992) used correlations, one-way analysis of variance (ANOVA), and multiple regression. Correlations of the MLAT were done with end-of-training ratings for speaking and reading proficiency (the FSI proficiency test is described above, under "Instrumentation") and with individual difference variables (see above for listing and descriptions of the instruments). The data used for the correlations between end-of-training proficiency and the MLAT Index for all language categories combined were filtered to equalize expected length of training and proficiency outcomes (that is, to make results of a language like French comparable to those of a language like Chinese).

Results

Distributions

Table 1 shows that the Index Score is somewhat higher for

Table 1
MLAT Descriptive Statistics for the Index Score

Category	*N*	Mean	SD	Range	Mode	Skewness	Kurtosis
All Students	343	63	10	21-80	70	-.973	1.392
Category 1	169	59	12	21-80	61,70	-.808	.625
Categories 2-3	120	66	8	45-80	70	-.462	-.171
Category 4	54	63	10	26-78	64	-.900	.770

Minimum possible Index: 20; maximum possible Index: 80. Category 1: Western European languages; Category 2: Swahili, Indonesian, Malay; Category 3: Eastern European and non-Western languages (except Category 4 languages); Category 4: Arabic, Chinese, Japanese, Korean.

Category 2, 3, and 4 languages than for Category 1 languages in central tendency and range (see "Sample" for definitions of these categories). The part scores follow the same pattern.

1393.pdf

The distributions, with their high central tendencies and reduced space below the ceiling for FSI students, reflect several forms of preselection. The first is that many students have self-selected for foreign affairs careers. Most of these went through their agency's selection process. This process has already probably eliminated some of the students least likely to score well on the MLAT. Second, the MLAT Index Score is used for selection of students in FSI's parent agency's personnel system, along with other evidence of likely learning, especially evidence of previous language learning success. (Such selection is authorized in the personnel regulations for the U.S. Department of State, where it is clearly stated that evidence of learning success overrides the MLAT.)

Selection is done in the State Department's personnel system especially for non-Western European languages, for which training to the "professional" proficiency level (S-3, R-3) takes 44-88 weeks. Relatively low MLAT students (Index below 55 for Category 3 or 60 for Category 4 languages) with no other evidence of success are normally sent to Western European languages, hence Category 1 is the group where we find a relatively large range of tested aptitude.

Preselection makes analysis of the MLAT's predictive value for Category 3 and 4 languages in this sample difficult. On the other hand, in view of the expense entailed by 44-week and 88-week intensive language training, assignments personnel understandably seek every indication of likely success or lack of it, without reference to the needs of the researcher.

Other results are described under two rubrics: findings related to prediction of language learning success and findings related to diagnosis and student counseling. The former are quantitative; the latter are qualitative.

Results Related to Prediction of Language Learning Success

Correlations

Correlation coefficients for MLAT Index, Total, and part scores with S- and R-ratings range in the 40s and 50s for the MLAT when a broad range of scores is available, comparable with coefficients found originally by Carroll (1990). The Index Score tends to show higher correlations with end-of-training proficiency ratings than do the part scores or the Total. Correlations for the Index Score are shown in Table 2. Category 1: Western European languages; Category 2: Swahili,

1399

Indonesian, Malay; Category 3: Eastern European and non-Western languages (except Category 4 languages); Category 4: Arabic, Chinese, Japanese, Korean. S-rating: speaking and interactive listening; R-rating: reading.

Table 2
Correlations of MLAT Index Score with End-of-Training Proficiency Ratings

Language(s)	r	S-rating	r	R-rating
All	.44	(N = 343)	.40	(N = 341)
Category 1	.52	(N = 169)	.55	(N = 168)
Category 2-3	.34	(N = 120)	.35	(N = 120)
Category 4	.47	(N = 54)	.34	(N = 53)

Correlations are weakest for Category 2 and 3 languages and strongest for Category 1 languages, where there is the greatest range, and the distribution of MLAT scores closely resembles a normal distribution. For Categories 1-3, correlations with reading and speaking are roughly the same. In Category 4 languages, they are stronger for speaking than for reading. This difference may be due to a smaller range in reading scores (they are much lower for beginners than in other languages), or possibly because the MLAT does not address abilities needed for reading languages that use Chinese or Chinese-type characters—three out of the four Category 4 languages.

Analysis of Variance

This investigation was done only for the entire sample, because the numbers of subjects were not sufficient for Category 2 and 3, or 4 languages separately. In a study of the extremely strong and weak students in the sample, the bottom 3 to 4 % were contrasted against all others and the top 5 to 6 % against all others. Extreme students were selected on a formula that combined length of training, relative difficulty of language by category, and end-of-training scores. There were fewer students at the low end because the very weakest may be withdrawn well before their scheduled end of training. In addition, the training staff strive to find ways to maximize each student's strengths, so that as many students as possible can achieve their training goals, which in most cases are S-3, R-3. More detail on the study of extremes,

1400

including the selection formula, is available in Ehrman (1994b).

Data for the individual difference variables were analyzed using the one-way analysis of variance procedure in SPSS for Windows 6.1. Levene's Test and t-tests for equality of means were used to determine unequal variances and the appropriate significance level, depending on whether variances were equal or unequal. The findings for the MLAT are displayed in Table 3.

Table 3
Performance Extremes: ANOVAs

Weakest, Speaking *N* selected (weakest): 4 (Parts & Total), 6 (Index)
N not selected (all others) = 292 (Parts & Total), 337 (Index).

Part	Weakest Mean	All Others Mean	Weakest SD	All Others SD	F	Sig.
1	24.5	36.5	6.5	9.1	6.8524	.0093
2	18.5	24.7	3.5	4.5	7.3634	.0070
3	11.0	28.3	8.6	9.9	12.1415	.0006
4	15.3	28.0	5.3	7.5	11.4289	.0008
5	11.5	19.3	4.7	5.3	11.4289	.0008
Total	80.8	136.7	24.6	27.5	16.3881	.0001
Index	43.2	62.7	10.8	10.5	20.5548	.0000

Strongest, Speaking *N* selected (strongest): 14 (Parts & Total), 19 (Index)
N not selected (all others) = 281 (Parts & Total), 324 (Index).

Part	Strongest Mean	All Others Mean	Strongest SD	All Others SD	F	Sig.
1	40.5	35.0	4.9	9.7	4.4395	.0362
2	27.1	24.3	2.8	4.7	5.2765	.0225
3	32.8	27.0	7.0	14.2	4.5701	.0336
4	30.0	27.2	5.0	7.9	1.7067	.1927
5	20.8	18.8	4.2	5.5	1.6950	.1942
Total	151.2	132.5	13.8	29.6	5.7291	.0175
Index	68.2	60.9	5.9	11.2	7.8286	.0055

1402

THIS PAGE INTENTIONALLY LEFT BLANK

Weakest, Reading *N* selected (weakest): 3 (Parts & Total), 4 (Index)
N not selected (all others) = 292 (Parts & Total), 337(Index).

Part	Weakest Mean	All Others Mean	Weakest SD	All Others SD	F	Sig.
1	23.0	36.4	7.0	9.1	6.4559	.0115
2	17.7	24.7	3.8	4.5	7.1481	.0079
3	7.3	28.2	5.5	9.9	13.4109	.0003
4	13.0	28.0	3.5	7.5	11.8901	.0006
5	11.0	19.3	5.6	5.3	7.3757	.0070
Total	72.0	136.6	21.2	27.6	16.3758	.0001
Index	40.5	62.7	12.6	10.5	17.6391	.0000

Strongest, Reading *N* selected (strongest): 78 (Parts & Total), 93 (Index)
N not selected (all others) = 217 (Parts & Total), 248 (Index).

Part	Strongest Mean	All Others Mean	Strongest SD	All Others SD	F	Sig.
1	38.9	33.8	6.3	10.5	15.0647	.0001
2	26.1	23.7	3.5	4.8	15.4653	.0001
3	31.0	26.9	8.6	10.2	14.7692	.0002
4	29.2	26.7	6.5	7.9	6.1293	.0140
5	21.3	17.9	4.1	5.6	22.5703	.0000
Total	146.5	128.0	20.9	30.1	23.7211	.0000
Index	66.3	59.6	8.0	11.3	26.1914	.0000

Data analysis done by SPSS for Windows v. 6.1, One Way Analysis of Variance Test. Degrees of freedom are available upon request.

Speaking. Of all the variables analyzed, Parts 3, 4, 5, the Total, and the Index scores best differentiated the weakest students. The MLAT variables also differentiated these weak students better than any other of the many variables in the research project.

For the strongest students' speaking scores, the Index ($F=7.83$, $p < .0055$) was the strongest differentiator from among the MLAT and learning style variables, but it was not as good as these biographical background variables: education level, number of previous languages,

1403

and previous highest score in speaking and especially reading (see Tables 4 and 5). The MLAT appears to differentiate the strongest speakers less clearly than the weakest speakers and the strongest and weakest readers.

Table 4
Results of ANOVAs for Weakest Students Compared with All Others

	Weakest Students			Non-Weakest Students					
Category	N	Mean	(SD)	N	Mean	(SD)	F	df	Sig.
From the Biographic Data									
No Prev.Lang.	15	1.0	(1.7)	674	1.7	(1.3)	7.1502	687	.008
Modern Language Aptitude Test									
Part 1	4	24.5	(6.5)	292	36.5	(9.1)	6.8524	294	.009
Part 2	4	18.5	(3.5)	292	24.7	(4.5)	7.3634	294	.007
Part 3	4	11.0	(8.6)	292	28.3	(9.9)	12.1415	294	.0006
Part 4	4	15.3	(5.3)	292	28.0	(7.5)	11.4289	294	.0008
Part 5	4	11.5	(4.7)	292	19.3	(5.3)	8.7868	294	.003
Total Score	4	80.8	(24.6)	292	136.7	(27.5)	16.3881	294	.0001
Index Score	6	43.2	(10.8)	339	62.7	(10.5)	20.5548	343	.0000
Hartmann Boundary Questionnaire *(Higher Scores Indicate Thinner Boundaries)*									
Total Score	3	187.7	(1.5)	165	246.9	(39.9)	6.5579	166	.01
Myers-Briggs Type Indicator *(Scores below 100 indicate sensing scores above 100 indicate intuition)**									
Sensing-Intuition	15	88.3	(29.0)	738	103.4	(29.9)	3.7513	751	.05
Learning Styles Profile									
Simultaneous Visual Processing	3	3.0	(0)	199	4.3	(1.1)	4.3550	200	.04

•Excerpted from Ehrman (1994b), Table 1.

1404

Table 5
Results of ANOVAs for Strongest Students Compared with All Others

	Strongest Students			Non-Strongest Students					
Category	N	Mean	(SD)	N	Mean	(SD)	F	df	Sig.
From the Biographic Data									
Education Level	27	4.1	(1.2)	645	3.3	(1.1)	13.3136	670	.0003
No Previous Languages	26	2.3	(0.7)	637	1.6	(1.0)	10.0750	661	.002
High Previous Speak Score	21	3.2	(1.0)	331	2.2	(1.1)	17.0908	350	.0000
High Previous Read Score	21	3.5	(0.9)	325	2.3	(1.1)	23.0790	344	.0000
Age	27	33.7	(7.9)	584	39.4	(9.3)	9.6396	609	.002
Modern Language Aptitude Test									
Part 1	14	40.5	(4.9)	224	35.0	(9.7)	4.4395	236	.04
Part 2	14	27.2	(2.8)	224	24.3	(4.7)	5.2765	236	.02
Part 3	14	32.8	(7.0)	224	26.9	(10.2)	4.5701	236	.03
Total Score	14	151.2	(13.8)	224	132.2	(29.6)	5.7291	236	.02
Index Score	19	68.2	(5.9)	269	60.9	(11.2)	7.8286	286	.006
Hartman Boundary Questionnaire *(Higher Scores Indicate Thinner Boundaries)*									
Tolerate Lack of Orderliness	153	23.7	(5.0)	9	19.5	(6.3)	93.8905	160	.05
Myers-Briggs Type Indicator (TDI) *(Scores below 5 indicate sensing or judging poles; scores above 5 indicate intuition or perceiving poles)*									
Realistic-Imaginative	21	7.2	(2.9)	575	5.7	(3.2)	4.5036	594	.03
Methodical-Emergent	21	4.6	(2.8)	575	3.2	(2.5)	6.0914	594	.01

•Excerpted from Ehrman (1994b), Table 2.

Reading. For reading, Parts 3 and 5 and the Total and Index Scores best differentiate the weakest students. The strongest are differentiated clearly by all MLAT parts except Part 4, with the Index Score providing the clearest distinction.

Multiple Regression

An exploratory stepwise multiple regression analysis for end-of-training speaking and reading examined the effects of age, education level, number of previous languages studied, highest previous speaking and reading ratings, a general motivation rating, two self-efficacy ratings (self-rated aptitude and expectation of success in this course), two anxiety ratings (for the course in general and about speaking in class), and the MLAT Index Score.

For speaking, the analysis yielded a multiple R of .40, R Square of .16, with two predictors in the equation: the MLAT Index Score (Beta .32, T = 3.293 p = .0014) and Highest Previous Reading Score (Beta .21, T = 2.208, p = .0297).

For reading, the analysis yielded a multiple R of .37, R Square of .14, with the same two predictors in the equation: the MLAT Index Score (Beta .27, T = 2.798, p = .0063) and Highest Previous Reading Score (Beta .22, T = 2.266, p = .0258).

Results Related to Diagnosis and Student Counseling

In this section, both quantitative and qualitative findings are described, as part of an ongoing effort to build learner profiles that can be used by teachers, teacher trainers, program managers, and even students themselves to enhance student learning. The quantitative results contribute to a fuller picture of the kinds of students who are advantaged and disadvantaged in full-time intensive and largely communicative language training, by adding personality factors to more cognitive abilities. The qualitative material is very exploratory, but it has been promising enough to merit description here so that others can use and test the emerging patterns. It is also included here because it provides more information on what the MLAT may actually be measuring, and because it sheds more light on the complexity of the apparently simple factor-analysis-based MLAT parts.

Relationships with Other Individual Difference Variables

There are other variables than the MLAT that are useful in the building of an individual learner profile that can be used for diagnosis and counseling (the utility of these for prediction is more directly addressed in Ehrman, 1993, 1994a, b; 1995b, 1996, Ehrman & Oxford, 1995; Oxford & Ehrman, 1995). These variables bear interesting relationships to the MLAT. Correlations of at least .30 between the MLAT Index Score and/or Total Score and other instruments used in the larger study are presented in Table 6. The correlations suggest the relationships described below.

Table 6
MLAT Index or Total Score Correlations with Other Variables

Variable	Lang. Category Group	rho	Correlate	N
Number of Previous Languages	All	.40**	Index	245
HBQ Prefer Blurred Edges	Cat. 1	.51*	Total	25
HBQ Prefer Low Neatness	Cat. 2	.47	Total	25
HBQ Thin External Boundaries	All	.32**	Total	102
HBQ Total Score (Thin)	All	.30**	Index	110
MBTI/TDI Intellectual (N)	Cat. 1	.45*	Index	96
MBTI/TDI Intellectual (N)	Cat. 2-3	.35**	Index	103
MBTI Intuition	Cat. 1	.34**	Total	93
MBTI Imaginative (N)	Cat. 1	.34**	Index	96
MBTI Introversion	Cat. 1	.30*	Total	93
LSP Simultaneous Processing	Cat. 1	.45	Index	24
LSP Sequential Processing	Cat. 1	.43	Index	24

All the above correlations are significant at least at the .05 level; * indicates the .01 level, ** indicates the .001 level. HBQ: Hartmann Boundary Questionaire, MBTI, LSP: Learning Style Profile. "Imaginative" and "Intellectual" represent intuitive poles of the MBTI/TDI Realistic-Imaginative and Pragmatic-Intellectual subscales for the sensing-intuition main scale.

Those who have scored high on the MLAT tend to have studied languages previously and often prefer an "intuitive" approach to taking in information on the MBTI. MBTI intuition indicates preferences for the abstract over the concrete, search for meaning, a preference for the "big picture" rather than details, and the speculative over the strictly experiential (Myers & McCaulley, 1985). They describe themselves as having relatively thin ego boundaries, especially with respect to such matters as dislike for too much neatness, order, and clear-cut separations among visual images. Thin ego boundaries, correlated with MBTI intuition, indicate receptivity to a wide range of experience, both internal and external, and a willingness to blur categories. This concept is used to operationalize a model of tolerance of ambiguity (Ehrman, 1993, 1996, 1998). High-MLAT students also are often more skilled at simultaneous and sequential visual processing on the Learning Style Profile (Keefe, Monk, et al., 1989).

The analyses of variance in the study of extremes support these findings for extremely strong and weak students and add as an advantage a preference for a flexible approach shown in the perceiving pole of one of the MBTI/TDI JP subscales, methodical versus emergent. (This subscale of the TDI scoring of the long MBTI opposes a desire to know in advance what will happen to a preference to let events "emerge" and cope with them as they come up; the strongest students indicated a preference for an emergent approach.) Some of the results from the study of extremes are displayed in Tables 4 and 5.

The MLAT and Learning Activities

A recent correlation study showed interesting relationships between the MLAT and a set of activities that students rated for perceived utility both before starting training and at the end of training (Ehrman, 1995). The results led to hypotheses about the meanings of high and low subscale scores in a given student's profile that are proving useful in student counseling, so that the MLAT can be used for more than gatekeeping.

The correlations were similar for both pre- and post-testing. Though the correlations were generally low (mostly 20s and some in the 30s), there seemed to be suggestive patterns in them when subjected to a content analysis. Findings described below were based on the content analysis of those items with which the MLAT was correlated (Table 7) and on correlations of MLAT scales with variables.

1408

Table 7
MLAT Index or Subscale Score Correlations with Items from Student Learning Activities Questionnaire

Pre-training questionnaire N=127; Post-training questionnaires N=various arranged from highest to lowest correlations

Items	Pre/Post Training	r	p	N
Index Score				
Positive				
Rates own ability as learner as good	Pre	.43	.000	127
Risk-taking is encouraged	Post	.40	.000	84
Forced to use what one knows to convey meaning	Post	.38	.000	86
Thinks has ability relative to FSI learners	Pre	.38	.000	127
Over-the-head reading is useful	Pre	.32	.000	127
Expectations of success in the course	Pre	.30	.001	127
Listen only to native speakers of target lang.	Post	.30	.000	79
Over-the-head listening is useful	Pre	.29	.001	127
Tries saying things one does not know the words for	Post	.26	.016	83
Study at home (vs. lab, library, other places)	Post	.25	.041	69
Role-plays, simulations, skits are useful	Post	.22	.050	79
Making presentations in the target lang. helps	Post	.22	.050	80
Negative				
Amt. of time spent making up tests for self	Post	-.55	.015	19
Amt. of time spent labeling pictures	Post	-.47	.041	19
Lang. Lab. (listening to tapes) useful	Post	-.33	.030	43
Amt. of time spent in lang. lab. listening to tapes	Post	-.33	.028	46
Go step-by-step so will not become confused	Pre	-.28	.002	127
Master one thing before going on to another	Post	-.26	.019	83

1409

Part 1: Number Learning

Positive

Risk-taking is encouraged	Post	.37	.001	74
Forced to use what one knows to convey meaning	Post	.25	.029	76
Likes to help design the program as it goes along	Pre	.20	.033	127

Negative

Amt. of time spent making up tests for self	Post	-.45	.051	19
Amt. of time spent studying notes	Post	-.31	.013	62
Master one thing before going on to another	Post	-.26	.028	74
Teacher's role is to make and execute a plan	Post	-.23	.050	74

Part 2: Phonetic Script

Positive

Tries saying things one does not know the words for	Post	.32	.005	74
Risk-taking is encouraged	Post	.30	.011	74
Forced to use what one knows to convey meaning	Post	.29	.011	76
Study at home (vs. lab, library, other places)	Post	.28	.031	62
Use authentic magazines and newspapers	Post	.27	.042	56
Listen only to native speakers of target lang.	Post	.27	.025	70
Finds pattern drilling useful	Post	.27	.022	72
Observing native speakers interacting in real conversation	Post	.26	.025	75
Thinks has ability relative to FSI learners	Pre	.25	.022	127
Wants to have a syllabus	Pre	.23	.001	127
Role-plays, simulations, skits are useful	Post	.23	.050	71
Rates own ability as learner as good	Pre	.21	.021	127
Prefers grammar explanations in the target lang.	Pre	.21	.021	127
Over-the-head reading is useful	Pre	.20	.036	127

1410

Negative

Amt. of time spent labeling pictures	Post	-.50	.029	19
Pronunciation not corrected unless unintelligible	Post	-.26	.028	73
Group study is part of the program	Post	-.25	.040	63
Go step-by-step not to confuse	Pre	-.21	025	127

Part 3: Spelling Clues

Positive

Using target language informally outside class (e.g., lunch)	Post	.37	.050	29
Listen only to native speakers of target lang.	Post	.28	.021	70
Oral transformation drills are useful	Pre	.27	.003	127
Thinks has ability relative to FSI learners	Pre	.25	.021	127
Correction focuses on formal features	Post	.24	.046	72
Rates own ability as learner as good	Pre	.22	.016	127
Listen to others speaking before trying to speak	Post	.22	.050	75

Negative

Making up tests for self is useful	Post	-.62	.011	16
Class content is related primarily to real life needs	Post	-.37	.001	73
Amt. of time in lang. lab. with tapes	Post	-.29	.040	43

Part 4: Words in Sentences

Positive

Listen only to native speakers of target lang.	Post	.44	.000	70
Using target language informally outside class (e.g., lunch)	Post	.36	.029	36
Forced to use what one knows to convey meaning	Post	.33	.003	76
Over-the-head listening is useful	Pre	.28	.002	127
Over-the-head reading is useful	Pre	.27	.005	127
Discover grammar patterns for self	Post	.26	.027	73
Risk-taking is encouraged	Post	.26	.026	74
Read without a dictionary	Post	.25	.035	70

Correction focuses on formal features	Post	.24	.043	72
Discover grammar patterns for self	Pre	.20	.032	127
Negative				
Amt. of time in lang. lab with tapes	Post	-.46	.002	43
Amt. of time spent labeling pictures	Post	-.36	.029	19
Lang. lab. (listening to tapes) is useful	Post	-.36	.027	39
Master grammar before using in communication	Post	-.32	.006	74
Go step-by-step so will not become confused	Pre	-.29	.002	127
Studying one's notes is useful	Post	-.29	.036	16
Group study is part of the program	Post	-.29	.022	63
Reading is limited to what student already knows	Post	-.25	.041	70
Teacher's role is to make and execute plan	Post	-.25	.029	74
Field trips are useful	Pre	-.24	.009	127
Teachers are flexible and alter plans as needed	Post	-.24	.040	74
Master one thing before going on to the next	Pre	-.20	.032	127

Part 5: Paired Associates

Positive				
Forced to use what one knows to convey meaning	Post	.29	.010	76
Listen to others speaking before trying to speak	Post	.25	.029	75
Over-the-head reading is useful	Pre	.23	.015	127
Negative				
Making up tests for self is useful	Post	-.65	.003	19
Recording self to listen to is useful	Post	-.49	.048	17
Proportion of time spent studying with other students	Post	-.41	.043	25
Lang. lab. (listening to tapes) is useful	Post	-.32	.048	39

In summary, high MLAT Index and the part scores correlate with items that are interpreted as reflecting self-confidence as a language learner and tolerance of ambiguity (low-structure activities and input). Very limited approaches, such as not moving on until something is mastered or strictly step-by-step learning, tend to correlate negatively with the MLAT.

The Index and Parts 2, 3, 4, and 5 are correlated with items suggesting acceptance of/preference for use of authentic material for reading and listening and authentic conversation.

Parts 3 and 4 are correlated with items suggesting endorsement of learning activities that reflect an analytic structured approach. This effect was slightly stronger for Part 3; students who rejected a "touchy feely" approach on one item (the only such item) also tended to be high scorers on Part 3.

In contrast, the Index and a strong peak score on Part 2 may suggest a more experiential approach, with many unstructured activities like role-plays that are supported by a syllabus and drilling so that learners do not have to do a lot of analysis on their own.

High scores on Part 3 and the Index appear to be related to comfort with unstructured auditory input.

Students who endorsed activities interpreted as indicating a preference for discovery learning tended to do well on the Index and Part 4.

Interpreting Part-Score Profiles

The above patterns suggested possible uses for the MLAT profile in student counseling, where they currently are being tested. Some profiles that these data suggest are outlined below.

1. All parts high (a very high Index will usually represent this kind of profile):
   - has done well on all the parts
   - self-confident as a learner
   - responds well to activities that require tolerance of ambiguity
   - likes relatively unstructured learning
   - enjoys and even prefers authentic input.

A related analysis found a relationship between endorsement of relatively unstructured, ambiguous, authentic activities and higher end-of-training scores (Ehrman, 1995a).

1413

2. A more uneven profile in which Parts 3 (especially) and 4 are high:

- analytic learner, perhaps field independent
- likes a program with a clear plan (not the same as a restrictively sequential program)
- usually has good knowledge of English vocabulary and grammar.

3. An uneven profile in which Part 2 is highest, together with a strong Index (most other parts above average), may indicate a student who likes experiential, hands-on, participatory learning and learns best from material in context.

4. An uneven profile in which Parts 2 and 4 are relatively high, together with a strong Index, may suggest a student who likes to take control of his or her own learning sequence and can use both analytic and global learning strategies comfortably.

5. When either Part 1 or Part 5 is the highest of the part scores, there so far seems to be little that is distinctive, though interviews are suggesting that low scores on Part 5 indicate either poor mnemonic skills or weak metacognitive strategies, or both.

6. All parts low (a very low Index will usually represent this kind of profile):

- has done poorly on all the parts
- often lacks self-confidence as a learner and subject to anxiety because of slow progress
- likely to be overwhelmed by unstructured and uncontrolled input
- will need a great deal of scaffolding for longer than most other students
- likely to progress slowly.

Overall Total score on the MLAT or the Index gives a useful crude measure when it is either very low or very high: a very low Total or Index score indicates weakness in all the factors; a very high score suggests strength in all the factors. When the Index falls in the middle range—roughly within a standard deviation of the mean—it becomes much more important to examine the "scatter" of the part scores.

Using Part Scores With Students

The student counseling activity uses the variations in part scores to initiate interpretations that are raised with the student to examine how he or she learns. Interpretation usually requires an interview of the student. Responses by students to the question "What happened when you were doing this part?" provide useful information about the skills tested in each part. Each of the MLAT factors probably represents a set of abilities. For example, Part 3 has proved particularly fruitful in the diagnostic process with students. Among the possible task requirements of this item are gestalt processing of the whole word; sound-symbol processing; rapid hypothesis testing of sound-symbol possibilities; shift in mental set; and semantic evaluation.

These task requirement possibilities are represented as student performance in the following six cases of poor outcome on Part 3, each of which is followed by implications for the classroom. The cases represent composites of responses actually received to the query about what happened while students were completing this sub-test. (Many examples of real cases with specific score profiles are in Ehrman, 1996.)

1. One student might have done poorly on Part 3 because of difficulty with the kinds of analytic activities often described as "field independent." This student is likely to have difficulty with induction of rules and patterns and with grammar-oriented activities that have little context. Students of this sort usually find more contextual learning helpful.

2. Another might do poorly on the same part because of a weak English vocabulary (among the possible causal factors: poor education, low intelligence). This student, if a native speaker of English,[7] may have difficulty with vocabulary learning (among other things) because a lack of concepts and background knowledge. The classroom may have to include activities to help this student build content background as well as language.

3. A third one experiences difficulties reorganizing schemata or with gestalt processing or shifting mental set. Part 3 makes considerable demands on a person's ability to shift mental set. Such a student may be more comfortable with relatively predictable activities and less so with open-ended ones and may need assistance in building skills for coping with the unfamiliar or unexpected.

4. Yet another student might have a *phonetic coding difficulty* of the sort described by Sparks, Ganschow et al. (1995), that is, working with sound-symbol relationships. He or she is likely to have corresponding low scores in Parts 1 and 2, which also require decoding of sounds. Such a student is likely to be handicapped in both speaking and reading and will need more time to absorb material. Kinesthetic input such as learning with realia, drawing, and acting things out is likely to help this student.

5. Links among extraversion, desire for language use outside the classroom, and MLAT Part 3 suggest a student with a *distractibility* factor. That is, a strongly extraverted student who is drawn to interpersonal interactions might not be as adept at the kind of focus that the puzzle solving aspect of Part 3 entails as one who tunes out the world more readily. Study strategies, including frequent breaks and setting up conditions to maximize concentration, might help a student who has difficulty concentrating.

6. Finally, a person who is reminded by Part 3 items of crossword puzzles and dislikes them has had an *affective* reaction which interferes with ability to use cognitive resources. Alternatives to "puzzle-solving" activities would probably help this sixth student, or perhaps cooperative learning when puzzle-like activities are part of the curriculum. The teacher would need to be alert to the affective impact of these activities.

Interpretation of a student's profile is made more complex by factors that can affect any or all of the parts of the test. In some cases, a low score on Part 3 (or any other part) may be the result of a mechanical error, such as marking in the wrong row of the answer sheet. Sometimes a student will say that he or she did not understand the instructions for a given part (this response raises questions about attention, motivation, or test-taking strategies). Some students ascribe low scores to fatigue, which is plausible especially for the later parts. Interpretation is further complicated by the fact that a student might suffer from several of these difficulties at once.

Discussion

Summary

Despite the effects of restricted range, skewed distribution, and relatively limited ceiling (because of negative skew for this high-end

196

sample), the MLAT remains the best predictor of the variables examined. In general, the Index Score is the most useful of the MLAT variables as a predictor (strong in all cases, and with highest correlation coefficients). Of the part scores, Part 3 is the strongest predictor. Part 3, with its dependence on knowledge of English vocabulary as well as ability to solve puzzles, may also be an indirect indicator of general intelligence. This would apply to both fluid ability, because of the cognitive restructuring required by the task, and to crystallized ability (vocabulary), and "g" or general intelligence, since general vocabulary is also considered to be the single best stand-in for overall intelligence (Anastasi, 1988, Wesche, Edwards, & Wells, 1982).

Is the MLAT more suitable for Western European languages than for non-Western languages? The question remains open. Correlations show stronger results for Category 1 languages than for 2, 3, and 4 languages. On the other hand, the substantial preselection of students suggested by the very skewed distribution and the restriction of range in the sample may account for this finding as much as appropriateness of the MLAT for non-European languages. Furthermore, the fact that the correlations for Category 4 language outcomes are actually better than those for Category 3 languages, despite substantial truncation of range, might suggest that the MLAT is actually a fairly strong predictor for these languages. (The higher correlations might also be related to the much smaller numbers for Category 4 languages.) We cannot test either hypothesis on the FSI language-student population as long as they are pre-selected using the MLAT.

Of the extended set of variables in the research project (including learning strategies, cognitive styles, motivation, anxiety, and personality variables), the MLAT Index Score also continues to be the strongest predictor, both in the correlation coefficients and ANOVAs of extremely weak and strong students. It is especially powerful as a selector of extremes.

In addition to the relatively crude information provided by the Index score that may help in selection for training, the part-score profile shows promise as a way to better target classroom interventions and advice to students about appropriate learning strategies to develop. Strong performance on the MLAT appears to be related to personality variables that indicate high tolerance for ambiguity and the ability to reconceptualize input (e.g., reanalyze, arrange hierarchically, find abstractions that reconcile apparent contradictions).

1417

Is the MLAT Passe in an Age of Communicative Teaching?

The MLAT has been criticized by many as rating aptitude only for audio-lingual training, which was in vogue when the MLAT was developed. However, the MLAT correlations remain about the same, although the teaching methodology has changed considerably (most FSI courses now have a substantial communicative component, and some are almost wholly communicative). Why is this so? The following are some possibilities.

1. Perhaps the MLAT is really multidimensional, and a different set of dimensions applies to different methodology.

2. Perhaps the operative factor is really some form of coping with ambiguity or coping with the unfamiliar.

3. Possibly, it is the "g" (general intelligence)-factor that is operative for FSI students. (Sasaki (1993) found a general cognition factor, which she describes as similar to "g," to account for 42% of the variance among Japanese college students studying English as a foreign language.)

4. The very nature of classroom training may make a difference. Although FSI classroom training requires the ability to cope with communicative activities and access global and inferential learning, it also makes heavy demands on analytic skills. These may become increasingly important at higher proficiency levels. This fact may be why Parts 3 and 4 together are the most predictive of extremes in achievement, together with the Index, which is more associated with predilection for the more open-ended learning that is also necessary for achieving high proficiency levels in FSI classrooms. The study of ego boundaries using the Hartmann Boundary Questionnaire (Ehrman, 1993) found a similar construct, "tolerance of ambiguity," to be essential to effective classroom learning at FSI. In this study, thin ego boundaries that let a student take in new data were not enough alone—students had to impose some sort of mental structure on their intake and at the same time stay open to the fact that their structures were hypothetical. Investigation now under way is examining the applicability of the field independence construct to these findings, further information on which is to be found in Ehrman (1996, 1997).

The Aptitude Concept

Expanding the aptitude concept is one of the subjects of an ongoing investigation of individual differences in language learning.

1418

The subject is discussed in greater detail in Ehrman, 1994b, 1995b, 1996.

Among the outcomes of the study is evidence for an expanded definition of aptitude that includes both cognitive aptitude (measured specifically for languages by the MLAT and more generally by cognitive aptitude tests) and personality factors that predispose a learner to cope with ambiguity and apparent chaos. These become especially important in the relatively unstructured learning setting of communicative teaching approaches. A nexus is emerging of the following characteristics that seem to be related to success in the demanding intensive FSI classroom:

- cognitive aptitude (may include ability to cope with the un familiar)
- non-linear, discovery learning
- orientation to meaning over form
- ability to cope with surprises (linguistic and pedagogical)
- openness to input and tolerance of ambiguity
- ability to sort input, analyze as appropriate, and organize into mental structures.

The last is almost certainly related in some way to the much-studied field independence construct (e.g., Brown, 1994; Chapelle & Green, 1992; Ehrman, 1996, 1997; Witkin & Goodenough, 1981). It may be that the MLAT provides a way to measure field independence through verbal activities, in contrast to the usual tests of ability to disembed geometric figures (e.g., Witkin, 1969). Such a measure might improve the value of the field independence construct for language learning.

Absence of the above-listed characteristics appears to disadvantage FSI learners, perhaps more than the presence of these variables advantages those learners (Ehrman, 1994a, b, 1995b, 1996).

There seems to be a kind of aptitude-personality nexus that consists of cognitive flexibility, tolerance of ambiguity (including ability to impose structure on input), and ability to make use of non-linear learning strategies that are well suited to work with input in forms aimed at native speakers of the target language.

The MLAT is the most powerful of the predictive variables used, even in programs that are very different from those in vogue when it was designed. It may be that the ability to manage unfamiliar and contradictory input leads both to success in communicative classrooms and to high scores on the MLAT. The MLAT may gain its

relative power because it requires the examinee to cope with the unfamiliar on tasks that at least partially simulate language learning tasks. In contrast, personality inventories ask about general life preferences, and strategy inventories do not address how the strategies are used but only whether the student is aware of using them. "Faking good" is nearly impossible on the MLAT, and malingering is vanishingly rare at FSI.

Although the MLAT provides strong information about classroom language learning ability, it is supplemented by personality variables. The significant correlations between the MLAT and the personality measures, though not strong (between .21 and .33), are consistent across personality questionnaire and MLAT subscales (Ehrman 1993, 1994a, b, 1995b). In all cases, MLAT scores are linked with variables that suggest tolerance for ambiguity.[8]

The links between the MLAT and personality variables suggest a role for the disposition to use one's cognitive resources in ways that go beneath the surface and that establish elaborated knowledge structures. Those who are open to new material, can tolerate contradictions, establish hypotheses to be tested, focus on meaning, and find ways to link the new with previous knowledge structures seem to have an advantage in managing the complex demands of language and culture learning. The weakest students appear to be overwhelmed by the chaos they encounter; the strongest meet it head on, and may even embrace it to a degree.

As of now, the answer to the question "Is the MLAT passe?" is: Probably not, though it has much the same limitations as a sole *predictor* of learning success that it has always had. It is quite good, especially if viewed as an indicator of learning dispositions that will affect classroom performance, but it probably should not be more than one tool in a toolkit. Scatter analysis of the part scores is a promising use for placement, counseling, and remediation, particularly in the hands of an evaluator who treats the scores as signposts to interpretations to be tested, not as absolute predictors.

Limitations of This Study

The greatest limitation of this study, like all those from FSI, is the question of generalizability. Use of a sample drawn from a high-end, pre-selected population in itself restricts range, affects distributions, and strongly indicates the need for replication with samples more typical of what the usual reader of this publication works with. For the MLAT, unlike any of the other instruments in the larger study, the use

1420

of the instrument itself to help pre-select the sample severely limits both the statistical normality of the sample and our ability to make inferences from the findings.

The impossibility of establishing a truly normal distribution of MLAT scores in this sample also means that the statistical tests that assume normal distributions and similar sample sizes are used in unconventional ways. The number of tests conducted increases the chance of type I errors (false positives), though the consistency of findings over a number of variables may reduce the likelihood of such error. For these reasons, the findings reported here must be considered suggestive, not conclusive.

Next Steps

There is much more to look at in these data in the course of trying to find out what the MLAT is good for and what are its limitations. Among these are to seek normally distributed samples on which to replicate this study, begin multiple regression and discriminant analysis to see if the MLAT is a better predictor in combination with other variables; and find out what has happened with subjects who return from overseas and are tested—are they improved, worse, the same?

On the qualitative front, continued investigation can seek to confirm the working hypotheses described above in the section on student counseling and systematize them for use by people other than researchers, so that the MLAT part scores can provide useful information about specific learning strengths and difficulties that can be used in curriculum design and interventions with individual students. Eventually, a quantitative study of the part-score profiles should be designed and undertaken.

Notes

[1]The remainder of the literature review owes much to a draft prepared by Frederick Jackson for an FSI roundtable at the Language Testing Research Colloquium in 1994 (Jackson, 1994).

[2]The MLAT Project is separate but overlaps with the Language Learning Profiles Project, especially because it uses the same data set.

[3]The Department of Defense uses a similar classification.

[4]Only three percent of students in this sample were studying

1421

Category 2 languages, a number which is too small for most analyses. Category 2 and 3 languages are therefore combined.

[5]Although Appendix A lists possible Index Scores below 20, currently used scoring devices do not yield Index Scores below 20.

[6]The MLAT was standardized in part on an FSI sample. Although that sample, as a result of the times (late 1950s) was all male, no gender differences have appeared on the MLAT among present students on any sub-test of the MLAT or on its Total or standardized score.

[7]The MLAT is designed for use with native speakers of English. At FSI it is considered invalid for non-native speakers, though if one takes it and does well (Index greater than 50), such performance is considered a promising sign. Low scores, on the other hand, are ignored.

[8]A very recent study (Ehrman & Leaver, 1997), also shows a correlation of the MLAT with self-report of 'field sensitivity' (Index, r=.58, Part 2 .61, Part 3 .46, all at a p level of 0001). Field sensitivity, discussed at greater length in Ehrman (1996, 1997), is the tendency or ability to absorb language *osmotically* from the surrounding environment, in contrast to the field independent emphasis or decontextualizing the context.

References

Adams, M. (1984). The Modern Language Aptitude Test (Percentile Ranks). Unpublished FSI document. Arlington, VA: Foreign Service Institute.

Anastasi, A. (1988). *Psychological testing*. 6th ed. New York: Macmillan.

Brown, H. D. (1994). *Principles of language learning and teaching* (3rd Ed.). Englewood Cliffs, NJ: Prentice Hall Regents.

Campbell, C. (1987). Survey of attitudes specific to the foreign language classroom. Unpublished manuscript.

Carroll, J. (1962). The prediction of success in intensive foreign language training. In Glaser, R. (Ed.), *Training research and education.* Pittsburgh, PA: University of Pittsburgh.

Carroll, J. (1990). Cognitive abilities and foreign language aptitude: Then and now. In T. Parry, & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11-29). Englewood Cliffs,

1422

NJ: Prentice Hall.

Carroll, J. & Sapon, S. M. (1959). *Modern Language Aptitude Test.* New York: Psychological Corporation.

Chapelle, C., & Green, P. (1992). Field independence/dependence in second language acquisition research. *Language Learning, 42,* 47-83.

Clark, J. (1986). *A study of the comparability of speaking proficiency across three government language training agencies.* Washington, DC: Center for Applied Linguistics.

Demuth, K. A., & Smith, N.B. (1987). The foreign language requirement: An alternative program. *Foreign Language Annals, 20,* 67-77.

Ehrman, M. E. (1993). Ego boundaries revisited: Toward a model of personality and learning. In J. E. Alatis (Ed.). *Strategic interaction and language acquisition: Theory, practice, and research.* Washington, DC: Georgetown University Press.

Ehrman, M. E. (1994a). The Type Differentiation Indicator and adult language learning success. *Journal of Psychological Type, 30,* 10 - 29.

Ehrman, M. E. (1994b). Weakest and strongest learners in intensive language training: A study of extremes. In C. Klee (Ed.) *Faces in a crowd: Individual learners in multisection programs.* Boston MA: Heinle & Heinle.

Ehrman, M. E. (1995a). Correlations between MLAT scales and preferred student learning activities. Unpublished data.

Ehrman, M. E. (1995b). Personality, language learning aptitude, and program structure. J. Alatis, (Ed.), *Linguistics and the education of second language teachers: Ethnolinguistic, psycholinguistic, and sociolinguistic aspects* (pp. 328-345). Washington, DC: Georgetown University Press.

Ehrman, M. E. (1996). *Understanding second language learning difficulties.* Thousand Oaks, CA: Sage Publications.

Ehrman, M. E. (1997). Field independence and field sensitivity. In Reid, J. (Ed.), *Understanding Learning Styles in the Second Language Classroom.* Englewood Cliffs, NJ: Prentice Hall Regents.

Ehrman, M. E. (1998). Ego boundaries and tolerance of ambiguity in second language learning. In J. Arnold (Ed.), *Affect in language learning.* New York: Cambridge.

Ehrman, M. E., & Christensen, L. (1994). *Language learner motivation and learning strategies questionnaire.* Unpublished

1423

manuscript.

Ehrman, M. E., & Jackson, F. H. (1992). Classroom activities survey. Unpublished manuscript.

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal, 1,* 67-89.

Ehrman, M. E., & Leaver, B. L. (1997, March). Sorting out global and analytic functions in second language learning. American Association for Applied Linguistics Annual Meeting, Orlando, FL.

Hart-Gonzalez, L. H., & Ehrman, M. E. (1992). Study Activities Questionnaire. Unpublished manuscript.

Hartmann, E. (1991). *Boundaries in the mind: A new psychology of personality.* New York: Basic Books.

Jackson, F. H. (1994). Language aptitude. Draft talking points prepared for FSI roundtable on the Modern Language Aptitude Test, annual meeting of the Language Testing Research Colloquium,Washington, DC.

Keefe, J. W., & Monk, J. S., (with Letteri, C.A., Languis, M., & Dunn, R). (1989). *Learning Style Profile.* Reston, VA: National Association of Secondary School Principals.

Lefrancois, J. & Sibiga, T. C. (May, 1986). Use of the Modern Language Aptitude Test (MLAT) as a diagnostic tool. Unpublished paper, no source. [Ontario Institute for Studies in Education?]

Lett, J. A., & O'Mara, F. E. (1990). Predictors of success in an intensive foreign language learning context: Correlates of language learning at the Defense Language Institute Foreign Language Center. In Parry, T., & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 222-260.) Englewood Cliffs, NJ: Prentice Hall.

Lett, J. A., & Thain, J. (1994). *The Defense Language Aptitude Battery: What is it and how well does it work?* Paper presented at the Language Aptitude Invitational Symposium, Arlington, VA.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists.

Norusis, M. J. (1992). *SPSS for Windows 5.0.1.* Chicago: SPSS.

Oxford, R. L., & Ehrman, M. E. (1995). Adults' language learning strategies in an intensive foreign language program in the United States. *System, 23,* 359-386.

Parry, T. S., & Child, J. R (1990). Preliminary investigation of the

1424

relationship between VORD, MLAT, and language proficiency. In Parry, T. & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 30-66). Englewood Cliffs, NJ: Prentice Hall.

Parry, T., & Stansfield, C. W. (1990). Introduction. In T. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp.1-10). Englewood Cliffs, NJ: Prentice Hall.

Parry, T., & Stansfield, C. W. (Eds.). (1990). *Language aptitude reconsidered.* Englewood Cliffs, NJ: Prentice Hall.

Petersen, C. R., & Al-Haik, A. R. (1976). The development of the Defense Language Aptitude Battery (DLAB). *Educational and psychological measurement, 6,* 369-380.

Pimsleur, P. (1966). *The Pimsleur language aptitude battery.* New York: Harcourt, Brace, Jovanovich.

Pimsleur, P. (1968, August). Aptitude testing. *Language Learning.* Special Issue, *3,* 73-78.

Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning, 43,* 313-344.

Saunders, D. (1989). *Type Differentiation Indicator Manual: A scoring system for Form J of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists.

Skehan, P. (1989). *Individual differences in second language learning.* London: Edward Arnold.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13*(2), 275-278.

Sparks, R. L., Ganschow, L., and Patton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native language and foreign language learning. *Journal of Educational Psychology, 87,* 638-655.

Spolsky, B. (1995). *Measured words: The development of objective language testing.* New York: Oxford.

Thain, J. (1992). DLAB II prototype development: Status report and CY92 plan. Technical report. Monterey, CA: Defense Language Institute Foreign Language Center.

Wesche, M. B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude.* Rowley, MA: Newbury House.

Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics,* 127-140.

1425

Wilds, C. P. (1965). MLAT Index Scores: Foreign Service Institute Curve-1965. Unpublished table. Arlington, VA: Foreign Service Institute.

Witkin, H. A., & Goodenough, D. R. (1981). *Cognitive styles: Essence and origins: Field dependence and field independence.* New York: International Universities.

Witkin, H. A. (1969) *Embedded Figures Test.* Palo Alto, CA: Consulting Psychologists.

Appendix A

MLAT Subscales

Part 1

Number Learning: This sub-test requires the examinee to learn four morphemes and interpret them in combinations that form numbers; it is entirely orally delivered. The sub-test is described in the Manual (Carroll & Sapon, 1959) as measuring part of memory and "auditory alertness" which play a part in auditory comprehension (showing how well one understands what one hears) of a foreign language.

Part 2

Phonetic Script: This sub-test requires the examinee to select a written equivalent (in Trager-Smith phonemic transcription) for an orally delivered stimulus. The MLAT Manual describes the sub-test as dealing with the ability to associate a sound with a particular symbol, as well as how well one can remember speech sounds. In addition, the sub-test is described as tending to correlate with the ability to mimic speech sounds and sound combinations in a foreign language.

Part 3

Spelling Clues: In this entirely written sub-test, an English word is presented in a very non-standard spelling. The examinee must select structure and thus expected to provide information about the ability to handle grammar in a foreign language. No grammatical terminology is used, so scores do not depend on specific memory for grammatical terms.

1496

Part 4

Words in Sentences: The stimulus is a sentence with a word or phrase highlighted. The examinee must indicate which part of another sentence matches the designated part. The sub-test is entirely in writing. It is described as dealing with the examinee's sensitivity to grammatical with their English equivalents and given some time to learn them. The words are then tested. This sub-test is said to measure the examinee's ability to memorize by rote—a useful skill in learning new vocabulary in a foreign language.

Part 5

Paired Associates: The examinee is presented with 24 foreign words he correct synonym. Vocabulary items are progressively more difficult, though the most difficult is probably within the repertoire of a college graduate. According to the Manual, scores on this part depend largely on how extensive a student's English vocabulary is. As in Part 2, it measures the ability to make sound-symbol associations but to a lesser degree.

Raw Score Total: Total of all five subscales.

Index Score: Originally a scaled (T) score used at FSI that is based on the Total. The original mean was 50, with a standard deviation of 10. These norms are now out of date; the Index is now simply a conversion of the raw Total into a scale ranging between 20 and 80. Local norms using the Index have not been formally established because the Index score using the original norms is deeply embedded in the agency's personnel system.

1427

Appendix B

Conversion Table for MLAT Raw Total and Index Scores

Raw Total	Index	Raw Total	Index	Raw Total	Index
0-9	15	67-68	37	125-127	59
10-12	16	69-71	38	128-129	60
13-15	17	72-74	39	130-132	61
16-18	18	75-76	40	133-135	62
19-21	19	77-79	41	136-137	63
22-23	20	80-82	42	138-140	64
24-26	21	83-84	43	141-143	65
27-29	22	85-87	44	144-145	66
30-31	23	88-90	45	146-148	67
32-34	24	91-92	46	149-150	68
35-37	25	93-95	47	151-153	69
38-39	26	96-97	48	154-156	70
40-42	27	98-100	49	157-158	71
43-44	28	101-103	50	159-161	72
45-47	29	104-105	51	162-164	73
48-50	30	106-108	52	165-166	74
51-52	31	109-111	53	167-169	75
53-55	32	112-113	54	170-172	76
56-58	33	114-116	55	173-174	77
59-60	34	117-119	56	175-177	78
61-63	35	120-121	57	178-180	79
64-66	36	122-124	58	181-182	80

Author

MADELINE EHRMAN, Foreign Service Institute U.S. Department of State, Mailing address: 1501 South Randolph St., Arlington, VA 22204-4012. Specializations: individual differences (learning styles and personality), interpersonal dynamics in second language education, and program leadership.

1428

Applied Language Learning
1998, Vol. 9, Nos. 1 & 2, pp. 71-105

Factors in the Prediction of Achievement and Proficiency in a Foreign Language

Richard L. Sparks
College of Mount St. Joseph

James Javorsky
Purdue University

Jon Patton and **Leonore Ganschow**
Miami University

This study examines the results of a factor analysis of a battery of native language and foreign language (FL) aptitude measures used to predict FL proficiency. The study involved two groups of high school FL learners completing a second year of FL study, sixty 10th and 11th grade females attending a private, single sex, college preparatory high school and a coeducational population of thirty-six 10th grade students in a public school. The two groups were combined to perform the analysis. The latent structure of the test battery was of interest to determine if the components that emerged in previous factor analyses to predict FL grade would differ from the components that emerged in the present study to predict oral and written FL proficiency. Three components emerged from the principal components analysis and were identified as: Verbal Memory, Phonological Coding/Recoding, and Cognitive Speed Plus. Results showed that the three components received similar loadings in the analysis; together, the three components contributed 63.28% of the variance in overall FL proficiency.

© 1998, Richard L. Sparks, James Javorsky, Jon Patton and Leonore Ganschow

1429

Recently, foreign language (FL) educators have developed guidelines to measure the extent to which students become proficient in the oral and written aspects of a FL. The American Council on the Teaching of Foreign Languages (ACTFL) defines proficiency as "what an individual can and cannot do (with a foreign language), regardless of where, when, or how the language has been learned or acquired" (ACTFL Proficiency Guidelines, 1989). In part, because FL proficiency tests require trained professionals and are time consuming, only a few studies have used these guidelines. These studies have shown that ratings on FL proficiency measures can be useful as criterion variables and are closely related to scores on norm-referenced tests of proficiency in similar language domains (see Ehrman & Oxford, 1995; Lett & O'Mara, 1990; Wilson, 1989; Wilson & Graves, 1995).

For over forty years, FL educators and researchers have investigated the best predictors of FL learning success. In the 1950s and 1960s, researchers developed FL aptitude tests. The underlying assumption of these tests was that FL aptitude is closely associated with the linguistic rule structures of language. Paul Pimsleur developed a measure of FL aptitude, the *Language Aptitude Battery* (LAB) (Pimsleur, 1966), that was closely associated with analysis of linguistic structures. The LAB is comprised of a Verbal Ability score (vocabulary, language analysis) and an Auditory Ability score (sound discrimination, sound-symbol association). John Carroll and Stanley Sapon published the Modern Language Aptitude Test (MLAT, Carroll & Sapon, 1959) based on Carroll's ground-breaking factor analytic studies, the results of which showed four independent variables to be important for FL learning: (1) phonetic coding; (2) grammatical sensitivity; (3) inductive language learning ability; and (4) rote memory (Carroll, 1962). In a comprehensive factor analytic study of cognitive abilities, Carroll (1993) proposed a three-stratum theory in which FL aptitude and FL proficiency are both distinct and unique factors. FL aptitude and FL proficiency appear in the first, or lowest, stratum of Carroll's hierarchical model of cognitive abilities. The second stratum of Carroll's theory encompasses eight unique abilities, two of which are called "crystallized intelligence" and "broad cognitive speed." Of importance to FL learning theory is that FL aptitude (including phonetic coding, grammatical sensitivity, spelling, and verbal (printed) language) is subsumed by "crystallized intelligence." In his Model of School Learning for the study of a FL, Carroll also proposed that time, i.e., "broad cognitive speed," is important for FL learning. (See Skehan, 1986, for a discussion of Carroll's Model of School Learning.)

1430

In a recent study in which they examined the correlational and predictive validity of a new language aptitude test similar to the MLAT, Parry and Stansfield (1990) found that the MLAT was the best overall instrument for predicting language learning success in an adult population. Lett and O'Mara (1990) found that another FL aptitude test, the Defense Language Aptitude Battery (DLAB) (Petersen & Al-Haik, 1976), along with high cognitive ability were consistent predictors of success among learners in intensive FL courses in more difficult FLs. Although hypothesized to offer potential in enhancing success in FL learning, non-cognitive variables (e.g., attitude, motivation, language learning strategies, personality, cognitive style) were not predictive of FL learning success. FL aptitude tests were widely used in the 1960s and 1970s, but then generally fell out of favor with FL educators and researchers.

Other researchers have speculated that the concept of FL aptitude needs to be reexamined because "FL aptitude tests such as the MLAT do not take into account new insights revealed by cognitive psychologists into the human learning process in general, and the language learning process in particular" (Parry & Stansfield, 1990: p. 2). In the 1970s and 1980s, FL educators began to emphasize the role that affective variables might play in successful FL learning. For example, Gardner and his colleagues studied the role of attitudes and motivation in FL learning and hypothesized that these two variables were relatively independent of language aptitude (e.g., see Garnder 1985, 1990; Gardner & Lambert, 1972). Other researchers have investigated anxiety in FL learning and speculated that there is a type of anxiety specific to FL learning (e.g., Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1991, 1994). Ehrman (1990) has used measures of personality (e.g., Myers-Briggs Type Indicator, or MBTI) to determine their importance in predicting FL learning. Qualitative findings of her research show that personality measures can provide information about the compatibility of teaching methods and teaching styles with the learning styles of individual learners; quantitatively, however, personality instruments appear to be weak predictors of language learning success (Ehrman, 1990; Ehrman & Oxford, 1995). Oxford (1990a, b) suggests that language learning strategies and a learner's cognitive style might be responsible for success or failure in learning a FL. She and her colleagues suggest that good language learners use more efficient language learning strategies and that FL learners can be trained to use strategies to improve FL learning.

Some researchers have challenged theories that emphasize a role for affective factors and language learning strategies in predicting

1431

FL success or failure. For example, Au (1988) has criticized Gardner's theories for failure to show a causal link between attitudes/motivation and FL learning. Oller (1981) has speculated that affective instruments may be unintentionally assessing language proficiency. Sparks and Ganschow (1991) support Oller's position, e.g., suggesting that most items on Horwitz's Foreign Language Classroom Anxiety Scale, a self-reported measure of anxiety, are related to expressive or receptive language and verbal memory skills, thus confounding the effects of anxiety and leading to speculation as to causal direction, i.e., which came first, the difficulties with language learning or the anxiety? They suggest that affective states, e.g., low motivation, high anxiety, are generally the consequence, rather than the cause, of good and poor FL learning. Tiedeman (1989) reviews cognitive styles research and finds that most measures of cognitive style are best interpreted as ability tests, not as measures of preferences for information processing. Skehan (1991) finds little evidence to suggest that style or strategy training improves FL learning and suggests that there is still the "worrying possibility that good [language] learners are ones for whom the use of effective strategies are possible, while for the poor language learners they are not" (p. 288). To the present authors' knowledge, previous research does not substantiate that affective variables, learning styles, or learning strategies are primary causal factors in FL learning, nor is research conclusive about whether these variables and others (e.g., beliefs about language learning, culture, gender) are predictive of FL proficiency.

Several years ago, Sparks, Ganschow, and Pohlman (1989) introduced into the learning disabilities and FL literature the hypothesis that the extent of one's proficiency in his/her native language is likely to have an impact on one's ability to learn a FL (see Sparks, 1995; Sparks & Ganschow, 1991, 1993a,b, 1995a). Their hypothesis is called the Linguistic Coding Differences Hypothesis (hereafter referred to as LCDH).[1] The term "linguistic coding" was coined initially by Vellutino and Scanlon (1986) to describe the language-based problems of children who had pronounced difficulties learning to read and spell their native language. Vellutino and Scanlon demonstrated that these children had difficulty primarily with the phonological/orthographic (sound and sound-symbol) and syntactic (grammar), but not the semantic (meaning) codes of language. In the LCDH, Sparks and his colleagues hypothesize that FL learning is built upon one's native language skills; that is, an individual's skill in the native language components — phonological/orthographic, syntactic, and semantic — serves as the foundation for successful FL learning. Further,

1432

they speculate that both native and FL learning depend on basic language learning mechanisms and that problems with one language skill, e.g., semantics, are likely to have a negative effect on both the native and FL systems.

Native language researchers also have shown that students with difficulties in reading and writing often have overt or subtle problems with their oral language skills (e.g., Catts, 1986; Crain, 1989; Mann, Shankweiler, & Smith, 1984). In their research, Sparks and Ganschow have hypothesized that the majority of students with FL learning problems are likely to display problems with the phonological/orthographic (and sometimes, syntactic) codes of language.

Since the introduction of the LCDH, its authors have conducted a number of research studies at the secondary and postsecondary levels of education. Results of these studies have provided strong empirical support for the hypothesis (Ganschow & Sparks, 1991, 1995, 1996; Ganschow, et al., 1991, 1994; Javorsky, Sparks, & Ganschow, 1992; Sparks, Artzer, et al., 1998; Sparks & Ganschow, 1993a,c, 1995b, 1996; Sparks, Ganschow, Artzer, & Patton, 1997; Sparks, Ganschow, & Javorsky, 1993; Sparks, Ganschow, & Patton, 1995; Sparks, Ganschow, & Pohlman, 1989; Sparks, et al., 1992, 1992a, b, 1996, 1997, 1998).

Sparks and Ganschow have posited that the concept of FL aptitude and its relationship to native language skills holds the most potential for better understanding: (a) why (and how) some students learn a FL better than others; and (b) why language aptitude measures are good predictors of FL learning and success. They have advocated the position that language variables are likely to contribute the largest part of the variance in successful FL learning (Sparks & Ganschow, 1995a). They have encouraged FL researchers to consider what is known in native language learning as the Assumption of Specificity (AOS) as a guide to future research in FL learning. (See Hall & Humphreys, 1982, and Stanovich, 1988 for a detailed description of the AOS). The concept of the AOS proposes that students with a particular learning problem have a cognitive deficit (or difference) that is reasonably specific to the task in question; with regard to FL learning, the deficit (or difference) is likely to be related to language because FL learning is a language-based task. Sparks and Ganschow suggest that researchers who wish to determine the underlying cognitive factors that may be related to good or poor FL learning should begin by carefully specifying the particular performance differences whose nature and origin are of interest. In their view, the performance differences between good and poor FL learners should be language-related because it is highly probable that differences in FL learning are related to the learning of

143

language. (See, also, Sparks, 1995, and Sparks and Ganschow, 1993b, for a discussion of the AOS). Thus, the LCDH focuses specifically on language variables in FL learning, much like Carroll's model of FL aptitude.

One method to test the soundness of a hypothesis is factor analysis, which identifies variables that correlate highly and those that do not. Ganschow, Sparks, Javorsky, and Patton (1992) conducted a factor analysis of a test battery used to test the native language and FL aptitude skills of not-at-risk and at-risk FL learners at the secondary level. In the study, they found three separable components which they labeled Phonology/Syntax, Cognition/Semantics, and FL Aptitude. The three components contributed approximately 60% of the variance in students' end-of-year FL grades. The Phonology/Syntax component had the most significant loading in the analysis and was comprised of measures of spelling, word recognition, and pseudoword reading as well as grammar. They interpreted this component to be highly related to phonology/orthography because the written language measures used in the test battery were primarily composed of phonological/orthographic items (e.g., spelling, pseudoword reading).

Sparks, Ganschow, and Patton (1995) conducted a factor analysis of a test battery used to predict end-of-year FL grade. The test battery was similar to the one used in the aforementioned study with a different group of FL learners completing the first year of a high school FL course. In that study, they found three components that they labeled Phonology/Orthography, Meaning, and FL Aptitude/Metalinguistic. Again, the Phonology/Orthography component had the most significant loading in the analysis and the three components contributed approximately 60% of the variance in the end-of-year FL grades of the students. In these studies the results showed that both lower level processing, the sound and sound/symbol system of language represented by the Phonology/Orthography component (see Koda, 1992), and higher level processing, the semantic and metalinguistic components of language represented by the Meaning and FL Aptitude/Metalinguistic components, contributed to performance in the FL classroom.

To the authors' knowledge, no factor analytic studies have been conducted on a battery of testing measures used to predict FL proficiency.[2] Of interest in this present study was the determination of components that might emerge in a factor analysis of a test battery to predict oral and written FL proficiency and whether the factors that emerged would be different from factors identified in two prior factor analyses that were conducted to predict end-of-year FL grades (see

Ganschow, et al., 1992; Sparks, Ganschow, & Patton, 1995). The purpose of the present study, then, was to determine the latent structure underlying a battery of native language and FL aptitude tests, most of which had been administered to the students prior to the first year of FL study. Also included were two measures obtained during the course of the study: end of first-year grades in the FL course and a FL word recognition measure administered at the end of the students' second year of FL study. The authors were also interested in determining which of the factors that emerged from the factor analysis were predictive of performance in English courses, FL courses, FL word recognition, and FL proficiency.

Method

Participants

Participants were 96 students enrolled in the second year of a FL course in the tenth and eleventh grades. Sixty students were females attending a highly selective, single sex, college preparatory high school and 36 students (17 males, 19 females) were attending a large, middle class, suburban public high school.[3] The mean age of the 96 participants was 16 years, 1 month (age range = 15 years, 5 months to 17 years, 4 months). The participants were enrolled in three FLs (Spanish = 52, French = 27, German = 17). All of the students had participated in the authors' previous studies (Sparks, Ganschow, & Patton, 1995; Sparks, et al., 1998). The justification for combining the data from two schools was that: (a) both groups had been administered similar test batteries; (b) the combined data sets resulted in the inclusion of a larger population; and (c) the combined data sets increased the probability that a more consistent factor pattern would emerge.

Instruments for Factor Analysis

Three types of instruments — native language, FL aptitude, and FL word recognition — were used to predict various aspects of FL proficiency and achievement, end of first-year FL grades, and eighth grade English grade. The predictor measures and outcome variables are described below.

435

Predictor Measures.

Measures common to both private and public school populations were used in the factor analysis: (a) the Wide Range Achievement Test-Revised (WRAT-R): Spelling Subtest; (b) Peabody Picture Vocabulary Test-Revised (PPVT-R); (c) Woodcock Reading Mastery Test-Revised (WRMT-R) Basic Skills Cluster: Word Identification and Word Attack subtests; and (d) the five subtests from the Modern Language Aptitude Test (MLAT) (Parts I-V).

Three measures in the analysis differed in the two populations, but were conceptually similar and measured similar skills. First, a group achievement test had been administered by both schools (High School Placement Test (HSPT): Total Test at the private school, Iowa Tests of Basic Skills (ITBS): Total Test at the public school). Second, a group-administered reading comprehension test had also been administered by both schools (Nelson-Denny at the private school, ITBS: Reading Comprehension subtest at the public school). Third, two different author-designed phonemic awareness measures were administered at both schools (Phoneme Deletion in the private school, Pig Latin in the public school). The HSPT and ITBS: Total Test were combined to form a Group Achievement variable; the Nelson-Denny and ITBS: Reading Comprehension subtest were combined to form a Reading Comprehension variable; and the Phoneme Deletion and Pig Latin measures were combined to form a Phonemic Awareness variable. For the phonemic awareness tasks, Z scores were computed by standardizing both variables and were used in subsequent analyses.

A list and description of testing instruments is presented in Appendix A. Abbreviations for each variable are used in Appendix A and in all Tables.

Outcome Variables

The outcome variables concerned the participants' performance in their native language and in the FL. Four types of outcome measures were collected: (a) end-of-year grades in the first year of FL high school instruction; (b) end-of-year grades in eighth grade English courses; (c) overall proficiency in reading, writing, and speaking/listening to a FL; and (d) FL word recognition. End-of-year FL and English grades were obtained from the participants' school records. In both subjects, grades represented scores on homework, projects, in-class activities, and oral and written quizzes and tests. The FL instructors reported that grades were comprised of approximately 25% listening, 25%

1436

speaking, and 50% reading and writing activities.

The outcome variables in FL performance involved the measurement of the participants' FL proficiency (i.e., skill in reading, writing, speaking, and listening to the FL). Three university-level FL professors, who were formally trained to administer proficiency tests in their respective languages of Spanish, French, and German, designed the tests according to guidelines developed by the ACTFL (1986, 1989). The overall proficiency score (identified as FL Total Test) included measures of written (reading and writing) and oral (listening and speaking) abilities in their respective FLs. Scores on the written and oral measures were combined to obtain an overall proficiency score.

Two measures were used to assess the student's proficiency in reading a foreign language. The first was a fictitious letter written in Spanish from Claudia Rivera, a high school student in Argentina, to a family in the United States. Claudia was planning to spend a year as an exchange student with the family to whom she was writing. The letter contained information about her, her family in Argentina, and a series of five questions that she wished to have answered prior to her arrival in the United States. The student was given 15 minutes to read the letter and answer ten multiple-choice questions in English about the contents of the letter. The second measure of reading proficiency was a slightly more difficult passage. The student was given 15 minutes to read a brief article from Selecciones (i.e., Readers Digest in Spanish) entitled "Los Palos de Punta" and answer ten multiple choice questions in English about the contents of the article. The student could achieve a combined maximum score of 20 on this reading comprehension measure.

To assess writing in a foreign language, the student was given 15 minutes to write a letter to answer Claudia's letter, incorporating the answers to Claudia's five questions in the response. ACTFL Guidelines were used in assigning a holistic proficiency level (i.e., one score based on all the criterion statements in a specific level of the ACTFL Guidelines) on the writing test. After the holistic score was determined, the student's performance was further defined for quantitative purposes by assigning a score of 0-5 on each of the following writing skills: vocabulary, cultural appropriateness, structures, comprehensibility, and spelling (0 = no production, 1 = Novice-Low, 2 = Novice-Mid, 3 = Novice-High, 4 = Intermediate-Low, 5 = Intermediate-High). A score of 0 was included in the scoring because some students at this level of education may have been unable to produce any response in Spanish. A student could achieve a maximum score of 25 on the writing measure.

To measure the student's ability to listen to and speak a FL, a 10-15 minute oral interview following ACTFL Oral Proficiency Interview guidelines was conducted individually with each student using the four phases prescribed in the ACTFL Guidelines: warm-up, level check, probes, and wind-up (Omaggio, 1986). The entire interview was audiotaped for later scoring. Prior to the beginning of the oral interview, the tester explained to the student in English that, after she had had an opportunity to chat for a few moments in Spanish, she would be given a conversation card in English to help her begin the conversation (Spinelli, 1988). The interview proceeded as a friendly conversation in Spanish about topics which naturally emerged as the student responded to the conversation and the interviewer guided the conversation through the phases listed above. The ACTFL Proficiency Guidelines for determining proficiency levels were used in assigning a holistic score (i.e., one score based on the ACTFL Guidelines) on the oral interview. After the holistic score was determined, the student's performance was further defined for quantitative purposes by assigning a score of 0-5 on each of the following skills: pronunciation, vocabulary, grammar, comprehensibility, and listening comprehension. The scoring procedure for the listening/speaking test was the same as in the FL writing test. A student could achieve a maximum score of 25 on the listening/speaking measure.

A student's total test score (FL Total Test) was the combination of her scores on the reading comprehension, writing, and listening/speaking tests. A student could achieve a maximum score of 70 on FL Total Test.

The reliability of the three proficiency subtests (reading comprehension, writing, listening/speaking) and the total proficiency test (FL Total Test) were checked by a Cronbach's Alpha calculation. For reading comprehension, the Cronbach's Alpha was .73; for writing, .76; for listening/speaking, .97; and for FL Total Test, .87.[4]

To assess word recognition in Spanish, students were also asked to read a list of 20 words in Spanish, some of which they had never seen before. This served to assess directly phonological/orthographic skills in Spanish, and indirectly, pronunciation ability. Each word was chosen because it contained a letter or letter combination with a phonetic sound (e.g., the /a/ sound in the Spanish word casa is different from the /a/ sound in the English word cat) or a phonetic element that is different in Spanish than it is in English (e.g., in the Spanish word teléfono, the primary stress fall on the second syllable). The vowel sounds in Spanish, diphthongs, words with diacritical marks, and multisyllabic words were included within the target words. Refer to

1438

Appendix B for sample lists of words in Spanish, French, and German. Due to an inadvertent error in communication, the German lists contained only half as many words as the Spanish and French lists, and the error was not noted until after completion of the testing. German students, however, were each administered two lists of ten words and, therefore, read the same number of words (20) as the Spanish and French students.

After the Spanish and French students' raw scores were obtained on the word recognition measure, raw scores were transformed into Z scores. For the German students who read two lists, raw scores were also transformed to Z scores. Each German student's score was the average of the standardized scores from the two lists. These standardized scores were used in all subsequent analyses.

Procedure

Eighth grade English grades and the results of the group achievement tests (HSPT and ITBS) were obtained by the authors from school records. The authors administered the native language measures and the FL aptitude test during the first quarter of the school year in which each participant was enrolled in a first-year FL course. The MLAT, WRAT-R: Spelling, and Nelson-Denny Reading Test were administered in groups. The PPVT-R, WRMT-R; Basic Skills Cluster, and the phoneme awareness measures (Pig Latin and Phoneme Deletion) were administered individually. Total test time was approximately two hours for the native language and FL aptitude measures.[5]

The oral (speaking/listening) and written (reading comprehension, writing) FL proficiency measures (FL Total Test) were administered at the end of the participants' second year of FL study. The FL reading comprehension test took ten minutes, the FL writing test took fifteen minutes, and the FL speaking/listening test took from ten to fifteen minutes to complete. The FL word recognition measure was administered after the FL speaking/listening test was completed. The interviewer randomly selected one of the five word lists in Spanish and French, each of which contained twenty words, and two of the five lists in German, each of which contained ten words. The word recognition measure took five minutes to administer. Both the FL Total Test and FL word recognition measure were administered at the end of the students' second year of FL study.

1439

Data Analysis

To determine the latent structure underlying the testing instruments, a factor analysis with varimax rotation was performed on the data. Because the goal of the factor analysis was to derive factor scores for use in a multiple regression analysis, a varimax rotation was selected (Tabachnick & Fiddell, 1996). The purpose of varimax rotation, which is the most commonly used in social science research, is to simplify factors by maximizing the variance of the loadings within each factor across variables. The spread in loadings is maximized — loadings that are high after extraction become higher after rotation and loadings that are low become lower. Thus, interpreting a factor is easier because it is obvious which variable(s) correlates with it. To select the number of factors to analyze, we used a minimum eigen value, which represents a variance of 1.0. Because the variance that each standardized variable contributes to an initial factor analysis extraction is 1.0, a component with an eigen value of less than 1.0 is not as important from a variance perspective as an observed variable (Tabachnick & Fiddell, 1996).[6]

The analysis examined the relationship among and between observed variables and their relationship to a set of unobserved components or theoretical constructs. This relationship appears as a factor loading that has a range of 0 to 1.0. The higher the loading of an observed variable, the more important the variable is to the factor, or construct. From the resulting factors, predictor scores were derived. In order to examine the predictive relationship between the predictor scores and the outcome variables, a regression procedure was applied.

To examine the predictive relationship between the factors and the outcome variables, a standard multiple regression analysis was performed using Factor scores, derived from the factor analysis, as dependent variables, and FL grades, English grades, FL word recognition, and overall FL proficiency as independent variables. Statistical significance for this analysis was set at $p \leq .05$.

Results

Three factors based on the 12 tests and subtests emerged from the factor analysis and accounted for 63.28% of the variance of the model. Table 1 shows the three components, the testing measures, their component loadings, and the communality estimates for each of the testing measures. Factor 1 (accounting for 21.58% of the variance) was defined as a Verbal Memory dimension. Factor 2 (21.55%) as a

1440

Table 1
Principal Components solution (with Varimax Rotation) and Communaliy Estimates on Test Battery in Combined Datasets

Test	Factor 1	Factor 2	Factor 3	Communality Estimates
MLAT I	**.71***	.19	.08	.55
MLAT V	**.71***	.22	.07	.55
PPVT-R	**.68***	.03	.25	.53
WRMT WATT	.17	**.81***	.32	.78
PHON AW	.21	**.77**	-.03	.64
WRAT-R SPELL	.13	**.69**	.47	.71
WRMT WID	.23	**.57***	.55	.69
MLATII	.51	**.53***	.13	.59
R COMP	.41	.06	**.76***	.75
MLAT III	-.03	.16	**.71***	.53
GROUP ACH	.57	.27	**.61***	.77
MLAT IV	.47	.26	**.51***	.55

Note. Boldface indicates the factor on which the test or subtest had the highest loading.

MLAT I = Modern Language Aptitude Test - Number Learning subtest;
MLAT V = Modern Language Aptitude Test - Paired Associates subtest;
PPVT-R = Peabody Picture Vocabulary Test-Revised;
WRMT WATT = Woodcock Reading Mastery Test-Revised (Word Attack subtest);
PHON AW = Phoneme Awareness;
WRAT-R SPELL = Wide Range Achievement Test-Revised (Spelling subtest);
WRMT WID = Woodcock Reading Mastery Test-Revised (Word Identification subtest);
MLAT II = Modern Language Aptitude Test - Phonetic Coding subtest;
R COMP = Reading Comprehension;
MLAT III = Modern Language Aptitude Test -Spelling Clues subtest;
GROUP ACH = Group Achievement; and
MLAT IV = Modern Language Aptitude Test - Words in Sentences subtest.

1441

Phonological Coding/Recoding dimension, and Factor 3 (20.16%) as a Cognitive Speed Plus dimension.

Factor 1 obtained substantial loadings (>.50) from 3 of the 12 tests and was labeled as defining the dimension Verbal Memory. This factor appeared to measure the cognitive dimension related to verbal memory and vocabulary for both native language and the FL and received heavy loadings from measures of verbal rote memory, oral language comprehension, and vocabulary involved in English and FL courses (Carroll, 1993).

Factor 2 obtained substantial loadings from 5 of the 12 tests and appeared to define the dimension Phonological Coding/Recoding. This factor appeared to measure the awareness that spoken words are composed of sounds, the ability to segment and manipulate sounds within words (both native and foreign), and the ability to decode written words (real words in the native language, FL words, nonsense words in the native language). This dimension addressed both the "meta" aspects of language at the sound level and skill in reading low frequency, unfamiliar, and difficult words to which the student has received little or no previous exposure. This dimension received heavy loadings from measures of phonemic awareness, FL phonology/orthography, and native language phonology/orthography.

Factor 3 obtained substantial loadings on 4 of the 12 tests and appeared to define the dimension Cognitive Speed Plus. This factor appeared to be related to both speed and language processing and production. Cognitive Speed may be described as "quickness in identifying elements, or distinguishing between elements, of a (visual) stimuli pattern, particularly when measured under pressure to maintain focused attention" (Horn, 1988, p. 666). Another aspect of Cognitive Speed involved the "quickness in deciding on answers"—a measure of "just how quickly one produces answers, both correct and incorrect, to problems of moderate difficulty" (Carroll, 1993, p. 615). In a review of 60 years of factor analytic studies, Carroll (1993) included the language processing dimensions of semantic processing speed (i.e., speed of verbal reception), word fluency (i.e., speed of word retrieval), verbal ability (i.e., general verbal knowledge), and semantic fluency (i.e., the speed of idea reception and production). Factor 3 was reflective of the language processing aspects of Cognitive Speed Plus and received heavy loadings from timed measures of FL syntax and native language phonological/orthographic processing, general vocabulary, and general knowledge (e.g., information in reading comprehension paragraphs).

A tolerance diagnostic was computed for each dependent variable and principal component to check for multicollinearity. The tolerance

1442

values, which ranged from .30 to .70, indicated that multicollinearity was not a threat in this data set (Tabachnick & Fidell, 1996).

Four multiple-regression analyses, which examined the relationship among the three Factor Scores and the four outcome variables, were performed with the following results.

FL Grades

In the prediction of first year FL grades, the multiple regression analysis yielded a solution that resulted in an R^2 of .381 and an adjusted R^2 of .361, $F(3, 92) = 18.87$, $p = .0001$. The regression indicated that all three Factor Scores, Verbal Memory, $F(1, 92) = 5.31$, $p = .0001$, Phonological Coding/Recoding, $F(1, 92) = 2.99$, $p = .004$, and Cognitive Speed Plus, $F(1, 92) = 4.41$, $p = .0001$, were significant in predicting first year FL grades (see Table 2).

English Grades

In the prediction of eighth grade English grades, the multiple regression analysis provided a solution that resulted in an R^2 of .289 and an adjusted R^2 of .266, $F(3, 92) = 12.45$, $p = .0001$. The regression indicated that all three Factor Scores, Verbal Memory, $F(1, 92) = 3.53$, $p = .0005$, Phonological Coding/Recoding, $F(1, 92) = 3.92$, $p = .0002$, and Cognitive Speed Plus, $F(1, 92) = 3.03$, $p = .0032$, were significant in predicting eighth grade English grades (see Table 3).

FL Word Recognition

In the prediction of FL word recognition, the multiple regression analysis yielded a solution that resulted in an R^2 of .353 and an adjusted R^2 of .331, $F(3, 92) = 16.67$, $p = .0001$. Two of the three regression coefficients, Verbal Memory, $F(1, 92) = 4.77$, $p = .0001$, and Phonological Coding/Recoding, $F(1, 92) = 5.11$, $p = .0001$, were significant in the prediction of FL word recognition proficiency (see Table 4).

Overall FL Proficiency

In the prediction of overall FL proficiency (FL Total Test), the multiple regression analysis yielded a solution that resulted in an R^2 of .201 and an adjusted R^2 of .175, $F(3, 92) = 7.73$, $p = .0001$. Two of the

1443

Table 2
Regression Model Using Factor Scores from Factor Analysis in the Prediction of First Year FL Grades

Factor	Standard Error	*B*	*t* value	*p*
Verbal Memory (Factor 1)	.06	.44	5.31	.0001
Phonological Coding/Recoding (Factor 2)	.06	.25	2.99	.0035
Cognitive Speed Plus (Factor 3)	.06	.36	4.41	.0001

Table 3
Regression Model Using Factor Scores from Factor Analysis in the Prediction of Eighth Grade English Grades

Factor	Standard Error	*B*	*t* value	*p*
Verbal Memory (Factor 1)	.053	.31	3.53	.0005
Phonological Coding/Recoding (Factor 2)	.053	.34	3.92	.0002
Cognitive Speed Plus (Factor 3)	.053	.27	3.03	.0032

1444

Table 4
Regression Model Using Factor Scores from Factor Analysis to Predict FL Word Recognition

Factor	Standard Error	*B*	*t* value	*p*
Verbal Memory (Factor 1)	.30	.39	4.77	.0001
Phonological Coding/Recoding (Factor 2)	.30	.43	5.11	.0001
Cognitive Speed Plus (Factor 3)	.30	.09	1.11	.269

Table 5
Regression Model Using Factor Scores from Factor Analysis to Predict Overall FL Proficiency

Factor	Standard Error	*B*	*t* value	*p*
Verbal Memory (Factor 1)	.72	.37	4.02	.0001
Phonological Coding/Recoding (Factor 2)	.72	.03	.38	.706
Cognitive Speed Plus (Factor 3)	.72	.25	2.63	.01

1445

three regression coefficients, Verbal Memory, $F(1, 92) = 4.02, p = .0001$, and Cognitive Speed Plus, $F(1, 92) = 2.63, p = .01$, were significant in the prediction of overall FL proficiency after two years of FL instruction (see Table 5).

Discussion

Results of the factor analysis on the combined data sets yielded three components. Together, the three factors accounted for a substantial percentage of the variance (63.28%) in FL proficiency. One of the factors, Verbal Memory (Factor 1), was represented by measures of oral language comprehension, vocabulary, and verbal rote memory. The authors speculated that Factor 1 was representative of the semantic (meaning) and verbal memory aspects of language. Both John Carroll's (1962) model of FL aptitude (i.e., phonetic coding, grammatical sensitivity, inductive language learning ability, rote memory) and Sparks and Ganschow's Linguistic Coding Differences Hypothesis (i.e., phonology/orthography, syntax, semantics) (Sparks, 1995; Sparks & Ganschow, 1991, 1993a,b, 1995a; Sparks, Ganschow, & Pohlman, 1989) emphasize that students' ability to use and understand language, generally, is likely to be the primary determinant of their ability to become proficient in the oral and written aspects of a FL. Whereas both Carroll's and Sparks and Ganschow's models of FL learning explicitly include semantics as necessary components of language proficiency, Carroll's model also includes rote memory.

All three factors appeared to measure conceptually different aspects of language proficiency, given their emergence as separate components in the factor analysis. Phonological Coding/Recoding (Factor 2) was represented by measures of phonemic awareness (the Phoneme Deletion and Pig Latin tasks), FL phonology/orthography (MLAT Phonetic Script subtest), and native language phonology/orthography (word recognition, pseudoword reading, spelling). Cognitive Speed Plus (Factor 3) was represented by *timed* measures of academic achievement (standardized group achievement tests), reading comprehension (standardized reading comprehension tests), and syntax (MLAT Words in Sentences subtest), and phonological/orthographic processing/vocabulary (MLAT Spelling Clues subtest). Although both Factors 2 and 3 were represented by measures of native language and FL skills, the authors speculated that they were separate components in the factor analysis because the testing measures in Factor 2, Phonological Coding/Recoding, all involved the coding and recoding of phonological/orthographic information, whereas the testing measures in

1446

Factor 3, Cognitive Speed Plus, all included a speed dimension.

Two of the measures in Factor 2, word recognition and spelling, are phonological coding tasks. Two of the other measures in Factor 2, pseudoword reading and MLAT Phonetic Script subtest, were hypothesized by the authors to be "phonological recoding" tasks. Phonological recoding is defined as "translating letters into sounds by application of letter-sound rules and then recognizing the identities of words from their pronunciations" (Ehri, 1992, p. 107). Gough (1984) suggests that phonological recoding is not used to read words to which the reader has had frequent exposure (e.g., word recognition in a student's native language). Instead, phonological recoding is used primarily to read nonsense words, low frequency words, unfamiliar words, and difficult words that have not received sufficient exposure to enter memory. Pseudoword (nonsense word) reading, then, was determined to be a phonological recoding task.

The MLAT Phonetic Script subtest was also hypothesized to meet the criteria for phonological recoding. On this task, students are asked to listen to the examiner (on a prerecorded tape) read aloud four words, each of which has a pronunciation that is different from the English sound/symbol system (e.g., *tik, tiyk, tis, tiys* are pronounced as /tik/, /tek/, /tis/, /tes/). After reading aloud five sets of four words each, the examiner returns to the first set and instructs the student to choose the one word (out of four) that is subsequently said aloud in each set. The Phonetic Script subtest includes six sets of five words each and introduces new sounds in each subsequent set. In a previous study that used a factor analysis procedure, the MLAT Phonetic Script subtest loaded on a separate component with other MLAT subtests (Sparks, Ganschow, & Patton, 1995). In another study a phonemic awareness measure, the Lindamood Auditory Conceptualization Test, loaded on a component that included the MLAT Phonetic Script subtest and three other MLAT subtests (Ganschow, et al., 1992). In the two aforementioned studies, pseudoword reading, as measured by the WRMT-R: Word Attack subtest, loaded on a native language phonology/orthography component. In the present study, phonological recoding tasks combined with a metaphonological task (phonemic awareness, i.e., Phoneme Deletion and Pig Latin) and emerged as a component important for FL learning.

In contrast to the Verbal Memory and Phonological Coding/ Recoding factors, two of the tasks in Cognitive Speed Plus (Factor 3), group achievement and reading comprehension, appeared to rely directly on students' processing and production of their native language under time constraints. The third testing measure in Cognitive Speed

144

Plus, the MLAT Spelling Clues subtest, relies directly on students' knowledge of English vocabulary and indirectly on the sound and sound-symbol system of English; this task is also administered under time constraints. On the Spelling Clues subtest, the student is presented with misspelled shortened versions of English words (e.g., luv, ernst, mblm, sidr). The instructions tell the student that the words are not spelled in the usual way, but are spelled approximately as they are pronounced; however, the words adhere to the English sound/symbol system although the students are not told this directly. The student must decode each word before finding a word from a list of five that corresponds most closely in meaning to the target word (e.g., student decodes *luv*, then chooses the word closest in meaning to *love* from the following words: *carry, exist, affection, wash, spy*). Previous studies suggest that success on the MLAT Spelling Clues subtest was likely to be equally dependent on students' knowledge of phonology/orthography as on their vocabulary knowledge (Ganschow, et al., 1992; Sparks, Ganschow, & Patton, 1995). The fourth measure in the Cognitive Speed Plus factor, MLAT Words in Sentences subtest, assessed grammatical sensitivity as well as the ability to process language under time constraints. Although not directly related to Carroll's model of FL learning, the factor Cognitive Speed Plus is similar to the "broad cognitive speed" factor in Carroll's (1993) three-stratum model of cognitive abilities. Findings show that the Cognitive Speed Plus factor appears to be related to those language tasks such as word retrieval and verbal fluency that are performed under time constraints. Moreover, Cognitive Speed Plus includes word retrieval and verbal fluency skills.

Results of the multiple regression analysis showed that all three factors were significant in predicting eighth grade English grade and end-of-year FL grades. This finding indicated that there were no differences in the factors predicting native language and FL classroom performance. The findings suggest that achievement in school-based FL language instruction relies, at least in part, upon the skills measured by the three factors.

In the prediction of FL word recognition, the results indicated that the Verbal Memory and Phonological Coding/Recoding factors were predictive of the ability to "crack the code" in order to read both native and FL words. One would have likely predicted that the several testing measures in Phonological Coding/Recoding (Factor 2), pseudoword reading, native language word recognition, spelling, phonemic awareness, and phonological/orthographic learning in a FL (i.e., MLAT Phonetic Script subtest), would be predictive of students' ability to read words in a FL. However, the authors had not anticipated

1448

that the tasks in Verbal Memory (Factor 1), verbal memory (MLAT Number Learning and MLAT Paired Associates) and receptive vocabulary (PPVT-R), would be predictive of the ability to read words in a FL.

One possible explanation for this finding is that students with stronger phonological/orthographic skills are those who read more frequently. Numerous researchers have found that exposure to print (i.e., reading) can predict a variety of behavioral outcomes even when stringent controls for background characteristics and general cognitive ability are utilized (see Anderson, Wilson, & Fielding, 1988; Guthrie, Schaefer, & Hutchinson, 1991; Stanovich & Cunningham, 1992). One positive behavioral outcome of more frequent exposure to print can be a stronger vocabulary (e.g., see Cunningham & Stanovich, 1991; Stanovich, 1993; West, Stanovich, & Mitchell, 1993). Thus, it is likely that students with strong phonological/orthographic skills (i.e., pseudoword reading, native language word recognition, spelling, phonemic awareness, FL word recognition) read larger amounts of text and subsequently acquire larger vocabularies as a result of reading. In the present study, then, recognition of FL words also might have been affected by the extent of students' exposure to print in their native language.

Results of the multiple regression analysis showed that two of the three factors, Verbal Memory (Factor 1) and Cognitive Speed Plus (Factor 3), were predictive of overall oral and written FL proficiency; Phonological Coding/Recoding (Factor 2) was not significant in the model.[7] We speculate that Phonological Coding/Recoding was not significant in predicting overall FL proficiency because native language word recognition and pseudoword reading as well as metaphonological skill (i.e., phonemic awareness) rely heavily on one's skills in phonological/orthographic processing. Research in native language reading has shown that word recognition and spelling skills are dependent on phonemic awareness and phonological/orthographic processing skills and that variation in phonological processing skill is the primary mechanism that determines early success in reading (see Brady & Shankweiler, 1991; Stanovich, 1992). These skills are acquired early in learning to read and spell (i.e., in kindergarten and first grade), especially for stronger readers who are more likely to be enrolled in FL courses in secondary school. Findings suggest that although reading and spelling FL words is important for FL learning, overall proficiency in a FL, at least for secondary level students (i.e., tenth and eleventh graders), may be more heavily dependent upon the skills measured by the Verbal Memory and Cognitive Speed Plus (i.e., language production and comprehension, vocabulary, verbal rote memory, speed of

1449

language processing, or language skills that can be measured by paper-pencil tasks).

An important finding of this study, is that Factor 2, Phonological Coding/Recoding, was found to be predictive of end of first-year FL *grades*, but not predictive of FL *proficiency* after two years of study. FL educators have speculated that FL grades do not necessarily reflect how well students can read, write, speak, and listen to the FL (i.e., proficiency). There are several reasons why their speculation may be accurate. One reason that FL grades but not FL proficiency would be predicted by phonological/orthographic skill is that educators sample more reading and writing in daily FL classroom work (e.g., reading the textbook, taking written tests and quizzes, spelling words) than is sampled on a FL proficiency test (i.e., listening and speaking in addition to reading and writing).

A second reason Phonological Coding/Recoding may be predictive of FL grades but not FL proficiency is that FL proficiency measures are more holistic in nature. That is, FL proficiency measures are developed and scored in such a way that they do not measure the discrete skills that FL teachers generally sample when assigning grades.

A third reason why the Phonological Coding/Recoding factor might be predictive of FL grades but not FL proficiency may be that the large majority of the study's participants generally had well-developed phonological/orthographic skills in both their native language and the FL. In another study involving the same 96 students (Sparks, Ganschow, et al., 1998), findings showed that their mean scores on the phonological/orthographic measures used in this study were in the higher end of the average range. Ehri (1985) hypothesizes that *written* language development positively enhances *oral* language development because written language serves as a "visual-spatial model for speech" and that acquisition of a written language system "works various changes on spoken language, particularly at the phonetic and lexical levels" (p.361). Thus, the secondary level students with well-developed phonological and orthographic skills may have been able to rely on their skills in, e.g., language processing and production, verbal rote memory, and vocabulary, because they did not have to struggle to read or spell new and unfamiliar FL words in the FL classroom. However, other students with significantly lower levels of phonological/orthographic processing skill may not be able to process and produce language in the FL or learn FL vocabulary words as easily as students with stronger phonological/orthographic processing skills because students with lower levels of phonological and orthographic processing skill have more difficulty with the phonological (i.e., phonemic

1450

awareness) and phonological and orthographic (e.g., word recognition, spelling) aspects of both their native language and the FL. This speculation is consistent with Sparks and Ganschow's Linguistic Coding Differences Hypothesis (Sparks, et al., 1995; Sparks & Ganschow, 1991, 1993b, 1995a). Also, the speculation is intuitively appealing because Sparks, Ganschow, and their colleagues have found that students with significantly lower scores on measures of phonology/orthography (e.g., pseudoword reading, spelling) achieve lower scores on FL proficiency measures than do students with higher scores on phonological/orthographic processing measures (see Ganschow, et al., 1997; Sparks, Artzer, et al., 1998; Sparks, Ganschow, et al., 1998).

Overall, the results of this factor analysis showed that all components of language (phonology/orthography, syntax, semantics, and verbal rote memory) are important for oral and written proficiency in a FL. New to the FL literature is the finding that speed of language processing (i.e., Cognitive Speed Plus) contributed to the variance not only in FL grades, but also in FL proficiency.

Implications

There are several implications of this study. First, students learning to read, write, speak, and listen to a FL are likely to rely on all components of language (i.e., phonological/ orthographic, syntactic, semantic, and verbal memory) to learn the new language. Heretofore, students' ability to use the phonological/orthographic component of language to read (and spell) new and unfamiliar words (i.e., the FL) has not always been considered as important as the other language components in FL instruction. Its importance is reflected in the finding that this component appeared to measure a conceptually different aspect of language proficiency and was a significant predictor of end-of-year FL grades. Students may benefit from FL instruction that emphasizes not only the meaning aspects of language, but also the new sound-symbol system of the FL (see Ganschow & Sparks, 1995; Sparks, Artzer, et al., 1998; Sparks, Ganschow, Artzer, & Patton, 1997).

Second, FL proficiency may rely not only on previous language learning (i.e., the components of one's native language), but also on the ability to learn unfamiliar components of a new language quickly and effectively, i.e., speed of language processing and production. In this study, speed proved to be important to students' FL grades and to their oral and written proficiency in a FL.

Third, the lack of a separable FL aptitude component (e.g., a component composed solely or primarily of MLAT subtests) in the

present study suggests that FL aptitude (i.e., potential to learn a FL) may be an extension of native language skills. This finding is consistent with and supportive of the Linguistic Coding Differences Hypothesis (Sparks, 1995; Sparks & Ganschow, 1991, 1993b, 1995a; Sparks, Ganschow, & Pohlman, 1989), which is conceptually similar to Carroll's (1962) model of FL aptitude embodied in the MLAT, and to Carroll's (1993) model in which he finds that FL aptitude (which was associated with phonetic coding, grammatical sensitivity, spelling, and verbal language) was subsumed by a "crystallized" intelligence factor. The finding is also supportive of Carroll's (1973) speculation that FL aptitude is a "residue" of native language skills.

Fourth, the findings suggest that emphasizing the written (i.e., reading, writing) aspects of a new language along with the oral (i.e., listening, speaking) aspects may be important for effective FL learning, especially with students who have histories of and/or current difficulty with recoding and spelling in their native language (i.e., phonology/orthography). Findings are supportive of Ehri's (1985) hypothesis that written language development positively enhances the development of oral language skills.

Finally, FL educators may want to further examine the relationship between grades assigned in classroom FL courses and oral and written proficiency in a FL. Recent research by Sparks, Ganschow, et al. (1998) has shown that students who achieve higher end-of-year FL grades (i.e., As and Bs) in first and second year FL courses have significantly stronger overall proficiency in a FL than students who achieve lower end-of-year FL grades (i.e., Cs, Ds, and Fs).

Some additional research questions that might be investigated are: Do grades in FL courses reflect skill in reading and writing the FL more than speaking/listening to the FL? Do grades in FL courses correlate more strongly with written or oral FL proficiency?

Notes

[1]The authors have previously changed the name of their hypothesis from "deficit" to "differences" to reflect the notion that FL learning skill occurs along a continuum from very good to very poor FL learners.

[2]In other investigations, the same participant sample was used (Sparks, et al., 1997, 1998). However, different analytic techniques, i.e., group comparisons and prediction of FL proficiency, not factor analysis, were employed in those studies.

1452

[3]We thank the following individuals for their patience and cooperation in the completion of this study: Joanne Brewer, Carol Dettenwanger, Marilyn Herring, Kim Icsman, Ann Johnson, Nancy Lamb, Ed Rudder, Mary Sies, Shirley Speaks, and Craig Ullery.

[4]Coefficient Alpha can be applied to both conventional tests where answers are marked "right" or "wrong" and to responses receiving different weights (McDaniel, 1994).

[5]We thank the following individuals for their assistance in administering and scoring the native language and FL aptitude tests: Sue Aielli, Loreli Albus, Diane Beck, Sue Jarvis, Jane Pohlman, Mikki Springer, Kim Stevens, Mary Thompson, and Connie Yoho.

[6]In the varimax rotation, there were a total of two negatively-loaded variables out of 36 variables. Both of these negative loadings were in the -.03 range and did not load significantly on the factors that were extracted.

[7]Although the adjusted R^2 resulted in a modest value of .175, this value does explain 17.5% of the variance in the model and provides useful information in the development of a theoretical model (Krathwohl, 1993).

References

American Council of the Teaching of Foreign Languages. (1986). *American Council of the Teaching of Foreign Language Proficiency Guidelines.* Hastings-on-Hudson, NY: Author.

American Council of the Teaching of Foreign Languages. (1989). *American Council of the Teaching of Foreign Language Proficiency Guidelines.* Hastings-on-Hudson, NY: Author.

Anderson, R., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23,* 285-303.

Au, S. (1988). A critical appraisal of Gardner's social-psychological theory of second language (L2) learning. *Language Learning, 38,* 75-100.

Brady, S., & Shankweiler, D. (Eds.). (1991). *Phonological processes in literacy.* Hillsdale, NJ: Erlbaum.

Brown, J., Bennett, J., & Hanna, G. (1981). *Nelson-Denny Reading Test.* Chicago, IL: Riverside.

Carroll, J. (1962). The prediction of success in intensive foreign

1453

language training. In T. Glaser (Ed.), *Training and research in education* (pp. 87-136). Pittsburgh, PA: University of Pittsburgh Press.

Carroll, J. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of Psycholinguistics, 2,* 5-14.

Carroll, J. (1993). *Human cognitive abilities: A survey of factor analysis studies.* NY: Cambridge University Press.

Carroll, J., & Sapon, S. (1959). *Modern Language Aptitude Test (MLAT): Manual.* San Antonio, TX: Psychological Corporation.

Catts, H. (1986). Speech production/phonological deficits in reading of disordered children. *Journal of Learning Disabilities, 19,* 504-508.

Crain, S. (1989). Why poor readers misunderstand spoken sentences. In D. Shankweiler & I. Liberman (Eds.), *Phonology and reading disability* (pp. 133-165). Ann Arbor, MI: University of Michigan Press.

Cunningham, A., & Stanovich, K. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83,* 264-274.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test-Revised.* Circle Pines, MN: American Guidance.

Ehri, L. (1985). Effects of printed language acquisition on speech. In D. Olson, N. Torrence, & A. Hildyard (Eds.), *Literacy, language and learning* (pp. 333-367). Cambridge, MA: Cambridge University Press.

Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ: Erlbaum.

Ehrman, M. (1990). The role of personality type in adult language learning: An ongoing investigation. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 126-178). Englewood Cliffs, NJ: Prentice-Hall.

Ehrman, M., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal, 76,* 67-89.

Ganschow, L., & Sparks, R. (1991). A screening instrument for foreign language learning problems: Evidence for a relationship between native and second language learning. *Foreign*

1454

Language Annals, 24, 383-398.

Ganschow, L., & Sparks, R. (1995). Effects of direct instruction in Spanish phonology on the native language skills and foreign language aptitude of at-risk foreign language learners. *Journal of Learning Disabilities, 28,* 107-120.

Ganschow, L., & Sparks, R. (1996). Foreign language anxiety among high school women. *Modern Language Journal, 80,* 199-212.

Ganschow, L., Sparks, R., Anderson, R., Javorsky, J., Skinner, S., & Patton, J. (1994). Differences in anxiety and language performance among high- and low-anxious college foreign language learners. *Modern Language Journal, 78,* 41-55.

Ganschow, L., Sparks, R., Javorsky, J., Pohlman, J., & Bishop-Marbury, A. (1991). Identifying native language difficulties among foreign language learners in college: A "foreign" language learning disability? *Journal of Learning Disabilities, 24,* 530-541.

Ganschow, L., Sparks, R., Patton, J., & Javorsky, J. (1992). Factors relating to learning a foreign language among high and low risk high school students and students with learning disabilities. *Applied Language Learning, 3,* 37-63.

Gardner, R. (1985). *Social psychology and second language learning.* London: Arnold.

Gardner, R. (1990). Attitudes, motivation, and personality as predictors of success in foreign language learning. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 179-221). Englewood Cliffs, NJ: Prentice-Hall.

Gardner, R., & Lambert, W. (1972). *Attitudes and motivation in second language learning.* Rowley, MA: Newbury House.

Gough, P. (1984). Word recognition. In P. Pearson, R. Barr, M. Kamil, & P. Rosenthal (Eds.). *Handbook of reading research* (pp. 225-253). New York: Longman.

Guthrie, J., Schaefer, W., & Hutchinson, S. (1991). Relations of document literacy and prose literacy to occupational and societal characteristics of young black and white adults. *Reading Research Quarterly, 26,* 30-48.

Hall, J., & Humphreys, M. (1982). Research on specific learning disabilities: Deficits and remediation. *Topics in Learning and Learning Disabilities, 2,* 68-78.

High School Placement Test. (1991). Bensenville, IL: Scholastic Testing Service.

Horn, J. (1988). Thinking about human abilities. In J. Nesselreade & R. Cattel (Eds.), *Handbook of multivariate experimental psychology* (pp. 645-685). New York: Plenum.

155

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70,* 125-132.

Iowa Tests of Basic Skills. (1990). Chicago, IL: Riverside.

Jastak, S., & Wilkinson, G. (1984). *Wide Range Achievement Test-Revised.* Wilmington, DE: Jastak Associates.

Javorsky, J., Sparks, R., & Ganschow, L. (1992). Perceptions of college students with and without specific learning disabilities about foreign language courses. *Learning Disabilities: Research and Practice, 7,* 31-44.

Koda, K. (1992). The effects of lower-level processing skills on foreign language reading performance: Implications for instruction. *Modern Language Journal, 76,* 502-512.

Krathwohl, D. (1993). *Methods of educational and social science research: An integrated approach.* White Plains, NY: Longman.

Lett, J., & O'Mara, F. (1990). Predictors of success in an intensive foreign language learning context: Correlates of language learning at the Defense Language Institute Foreign Language Center. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 222-260). Englewood Cliffs, NJ: Prentice-Hall.

MacIntyre, P., & Gardner, R. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41,* 85-117.

MacIntyre, P., & Gardner, R. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44,* 283-305.

Mann, V., Shankweiler, P., & Smith, S. (1984). The association between comprehension of spontaneous sentences and early reading ability: The role of phonemic representation. *Journal of Child Language, 11,* 627-643.

McDaniel, E. (1994). *Understanding education measurement.* Madison, WI: WCB Brown & Benchmark.

Oller, J. (1981). Research on the measurement of affective variables: Some remaining questions. In R. Andersen (Ed.), *New dimensions in second language acquisition research* (pp. 14-32). Rowley, MA: Newbury House.

Omaggio, A. (1986). *Teaching language in context: Proficiency-oriented instruction.* Boston, MA: Heinle.

Oxford, R. (1990a). Styles, strategies, and aptitude: Connections for language learning. In T. Parry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 67-125). Englewood Cliffs, NJ: Prentice-Hall.

1456

Oxford, R. (1990b). *Language learning strategies: What every teacher should know*. Boston, MA: Heinle.

Parry, T. & Standsfield, C. (Eds.) (1990). *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice-Hall.

Petersen, C., & Al-Haik, A. (1976). The development of the Defense Language Aptitude Battery (DLAB). *Educational and Psychological Measurement, 36*, 369-380.

Pimsleur, P. (1966). *Language Aptitude Battery*. New York: Harcourt, Brace.

Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing, 3*, 188-221.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13*, 275-298.

Sparks, R. (1995). Examining the linguistic coding deficit hypothesis to explain individual differences in foreign language learning. *Annals of Dyslexia, 45*, 187-214.

Sparks, R., Artzer, M., Patton, J., Ganschow, L., Miller, K., Hordubay, D., & Walsh, G. (1998). Benefits of multisensory language instruction for at-risk learners: A comparison study of high school Spanish students. *Annals of Dyslexia*, 48, 239-270.

Sparks, R., & Ganschow, L. (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *Modern Language Journal, 75*, 3-16.

Sparks, R., & Ganschow, L. (1993a). The impact of native language learning problems on foreign language learning: Case study illustrations of the linguistic coding deficit hypothesis. *Modern Language Journal, 77*, 58-74.

Sparks, R., & Ganschow, L. (1993b). Searching for the cognitive locus of foreign language learning problems: Linking first and second language learning. *Modern Language Journal, 77*, 289-302.

Sparks, R., & Ganschow, L. (1993c). The effects of a multisensory structured language approach on the native language and foreign language aptitude skills of at-risk learners: A follow-up and replication study. *Annals of Dyslexia, 43*, 194-216.

Sparks, R., & Ganschow, L. (1995a). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *Modern Language Journal, 79*, 235-244.

Sparks, R., & Ganschow, L. (1995b). Parent perceptions in the screening for performance in foreign language courses. *Foreign Language Annals, 28*, 371-391.

Sparks, R., & Ganschow, L. (1996). Teachers' perceptions of students'

1457

foreign language academic skills and affective characteristics. *Journal of Educational Research, 89,* 172-185.

Sparks, R., Ganschow, L., Artzer, M., & Patton, J. (1997). Foreign language proficiency of at-risk and not-at-risk learners over two years of foreign language instruction. *Journal of Learning Disabilities, 30,* 92-98.

Sparks, R., Ganschow, L., Artzer, M., Siebenhar, D., Plageman, M., & Patton, J. (1998). Differences in native language skills, foreign language aptitude, and foreign language grades among high, average, and low proficiency foreign language learners: Two studies. *Language Testing, 15,* 181-216.

Sparks, R., Ganschow, L., Fluharty, K., & Little, S. (1996). An exploratory study on the effects of Latin on the native language skills and foreign language aptitude of students with and without learning disabilities. *Classical Journal, 91,* 165-184.

Sparks, R., Ganschow, L., & Javorsky, J. (1993). Perceptions of high and low-risk students and students with learning disabilities about high school foreign language courses. *Foreign Language Annals, 26,* 491-510.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992a). Test comparisons among students identified as high-risk, low-risk, and learning disabled in high school foreign language courses. *Modern Language Journal, 76,* 142-159.

Sparks, R., Ganschow, L., Javorsky, J., Pohlman, J., & Patton, J. (1992b). Identifying native language deficits in high- and low-risk foreign language learners in high school. *Foreign Language Annals, 25,* 403-418.

Sparks, R., Ganschow, L., & Patton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology, 87,* 638-655.

Sparks, R., Ganschow, L., Patton, J., Artzer, M., Siebenhar, D., & Plageman, M. (1997). Prediction of foreign language proficiency. *Journal of Educational Psychology, 89,* 549-561.

Sparks, R., Ganschow, L., & Pohlman, J. (1989). Linguistic coding deficits in foreign language learners. *Annals of Dyslexia, 39,* 179-195.

Sparks, R., Ganschow, L., Pohlman, J., Artzer, M., & Skinner, S. (1992). The effects of a multisensory, structured language approach on the native and foreign language aptitude skills of high risk, foreign language learners. *Annals of Dyslexia, 42,* 25-53.

Spinelli, E. (1988). *Situation cards for oral evaluation: First year*

153

French, German, and Spanish. New York: Holt, Rinehart, & Winston.

Stanovich, K. (1988). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia, 38*, 154-177.

Stanovich, K. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 307-342). Hillsdale, NJ: Erlbaum.

Stanovich, K. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. In H. Reese (Ed.), *Advances in child development and behavior, 24*, 133-180. San Diego, CA: Academic Press.

Stanovich, K., & Cunningham, A. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory and Cognition, 20*, 51-68.

Tabachnick, B., & Fidell, L. (1996). *Using multivariate statistics*. New York: Harper Collins.

Tiedemann, J. (1989). Measures of cognitive styles: A critical review. *Educational Psychologist, 24*, 261-275.

Vellutino, F., & Scanlon, D. (1986). Linguistic coding deficits and metalinguistic awareness: Their relationship to verbal and code acquisition in poor and normal readers. In D. Yaden & S. Templeton (Eds.), *Metalinguistic awareness and beginning literacy* (pp. 115-141). Portsmouth, NH: Heinman.

West, R., Stanovich, K., & Mitchell, H. (1993). Reading in the real world and its correlates. *Reading Research Quarterly, 28*, 35-50.

Wilson, K. (1989). Enhancing the interpretation of a norm-referenced second-language test through criterion referencing: A research assessment of experience in the TOEIC testing context. Princeton, NJ: Educational Testing Service.

Wilson, K., & Graves, K. (1995). An assessment of validity-related properties of the secondary level English proficiency test and measures of ESL speaking and writing abilities in the Temple University-Japan context. Unpublished manuscript.

Woodcock, R. (1987). *Woodcock Reading Mastery Test-Revised*. Circle Pines, MN: American Guidance.

1459

Appendix A

Alphabetical List and Descriptions of Testing Instruments

High School Placement Test (HSPT TOT):

This group-administered, comprehensive placement test is designed specifically to aid in the selection and/or placement of students entering high school. The test measures Cognitive Skills (Verbal, Quantitative) and Basic Skills (Reading, Math, Language).

Iowa Tests of Basic Skills, Form J, Level 14 (ITBS TOT):

This test is a standardized measure of comprehensive growth in fundamental academic skills. It consists of subtests of language, reading, vocabulary, and mathematics, with questions presented in a multiple-choice format.

Iowa Tests of Basic Skills, Form H, Level 14 - Reading Comprehesion (ITBS RCOMP):

This test is comprised of paragraphs of varying lengths which the student reads and then answers questions presented in a multiple-choice format.

Modern Language Aptitude Test (MLAT):

This test measures foreign language aptitude using a simulated format to provide an indication of probable degree of success in learning a foreign language; it includes five subtests. The Long Form (MLAT LF) includes all five subtests. The subtests are: *MLAT I: (Number Learning)*: student learns numbers of a made-up language, and then transcribes spoken number words into written digits on hearing them presented rapidly; *MLAT II: (Phonetic Script)*: student listens to a sequence of syllables (many with no meaning in English) while looking at their graphemic transcriptions and is asked to quickly learn how the sounds (phonemes) correspond to the letters (graphemes); *MLAT III: (Spelling Clues)*: student reads English words presented as abbreviated spelling (e.g., luv) and then chooses the one word (out of five) that corresponds most nearly in meaning (e.g., carry, exist, affection, wash, spy); *MLAT IV: (Words in Sentences)*: student reads a "key" sentence in which

1460

a word is underlined, reads another sentence in which five words and phrases are marked as possible choices, and chooses the word or phrase in the second sentence that has the same grammatical function as the marked word or phrase has in the "key" sentence; and *MLAT V: (Paired Associates)*: student memorizes a list of nonsense words with their assigned English meanings.

Nelson-Denny Reading Test (NELSON), Form E:
This test consists of a series of eight paragraphs that measure the ability to read and answer multiple-choice comprehension questions in a timed format.

Peabody Picture Vocabulary Test-Revised (PPVT-R), Form L:
This test measures receptive vocabulary for Standard American English.

Phoneme Deletion (Phoneme Deletion):
This informal phonemic awareness measure has twenty items that test the ability to delete an initial, final, or medial phoneme and form a spoken word.

Pig Latin (Pig Latin):
This informal phonemic awareness measure has fifteen items that test the ability to delete the initial phoneme from a spoken word, move the phoneme to the end of the word, and then add an /a/ sound to the end of the new word. For example, the student must say "lackba" for *black*. The measure is composed of one, two, and three-syllable words.

Wide Range Achievement Test-Revised (WRAT-R SPELL), Spelling subtest:
This test measures performance on writing single words from dictation.

Woodcock Reading Mastery Test-Revised (WRMT BSK), Form G:
The *Basic Skills Cluster* tests two aspects of reading: *Word Identification* (WRMT WID) tests ability to read isolated words (e.g., urgent, hysterical, causation, heterogeneous); and *Word Attack* (WRMT WATT) tests ability to read (pseudo) nonsense words (e.g., dee, poe, vunhip, mancingful).

1461

Appendix B

Sample Word Recognition Lists for Spanish, French, and German

Spanish		French	
List 1	List 2	List 1	List 2
anoche	agosto	adorable	difficile
enero	efecto	midi	chimie
isla	ideal	robe	bonne
orilla	oriente	haise	caisse
usted	urbano	difficile	adorable
mesa	linda	croix	soie
señora	compañero	quest	coûte
jefa	junta	soeur	jeune
entrenamiento	historiadora	magnifique	peigner
salon	peatón	aéroport	aéroport
inventado	invierno	soleil	sommell
mural	musa	Joël	coïncidence
la amada	la alemana	longtemps	contravention
agencia	agitar	pluie	fruit
filólogo	préstamo	médicament	médicament
alto	algo	rue	vue
corriente	cerradura	hier	mariage
antena	anterior	décidé	télé
repentinamente	cuidadosamente	nationalité	nationalité
extensión	preocupación	réveillon	baigner

German	
List 1	List 2
Zehn	oval
Oktav	Bergkundige
Sieh	Zürich
Jawohl	Baümen
Jodeln	Juni
Deshalb	Tastatur
Postboten	Notiz
Büro	entgegennehmen
Imperfekt	Geschwindigkeit
Deutsch	Begrenzung

1462

Authors

RICHARD L. SPARKS, Associate Professor of Special Education, College of Mt. St. Joseph, Cincinnati, Ohio. *Specializations:* foreign language learning, foreign language learning problems, and reading disabilities.

JAMES JAVORSKY, doctoral candidate in Special Education, Purdue University, West Lafayette, Indiana. *Specializations:* attention deficit disorder, reading disabilities, and foreign language learning problems.

LEANORE GANSCHOW, Professor Emerita of Special Education, Miami University, Oxford, Ohio. *Specializations:* language disabilities (native, foreign, oral, and written).

JON PATTON, Visiting Professor and Senior Research Consultant, Miami University, Oxford, Ohio. *Specializations:* Statistical modeling and consulting about statistical packages.

1463

Applied Language Learning
1998, Vol. 9, Nos. 1& 2, pp. 107-119

Essay Scores as Instruments for Placement and Advancement in an Intensive English Program

Lynne Davis
Ruth Johnson
Floyd Olive
Southern Illinois University

Can essay scores of English-as-a-second-language (ESL) students be used as reliably for placement and/or advancement purposes as their Test of English as a Foreign Language (TOEFL) scores? Can these essay scores be used as reliably for placement and/or advancement purposes as recommendations for placement and/or advancement made by classroom teachers? What is the relationship between Test of Written English (TWE) sample essay scores and TOEFL scores when they are used to make decisions about ESL students' progression from one level to another within an intensive English program (IEP)? And what is the relationship between TWE sample essay scores and teacher recommendations when these are used to make placement and/or advancement decisions?

Exams such as the TOEFL that are usually used for placement, including placement into writing classes, are advantageous because they provide quick evaluation and objectivity (Leki, 1991; Perkins, 1983; Perkins & Pharis, 1977). Those who argue in favor of using such indirect measures of writing ability, such as measures of grammar, usage, word choice and syntax, say such tests actually measure the construct also measured by essay exams (in this case, English-language writing ability); that is, objective tests of writing have construct validity (Educational Testing Service, 1992). They also maintain that an objective, indirect measure of writing ability is superior to an essay test; it prevents the test taker from using avoidance strategy to evade using constructions he or she does not know or is unsure of using (Diederich, 1974).

© *1998, Lynne Davis, Ruth Johnson, and Floyd Olive*

1464

Having noted these supporting arguments for such objective tests, it is important also to note that objective tests have been criticized for not testing skills necessary for classroom performance (Leki, 1991). Critics argue in favor of an essay exam, saying that a writing test is the appropriate instrument for measuring writing ability/skill (Kroll, 1991; White, 1993). In addition, they say that essay scores reveal not only writing ability, but also "a robust measure of global language proficiency" (Kaczmarek, 1977, p. 159). This finding, says Leki (1991), is crucial when testing ESL students because it is not only their writing ability that must be determined, but also their general proficiency in English in order to place them appropriately into writing classes.

It should be noted here regarding the holistic grading method that recent discussions question the value of holistic scoring, as used in the TWE, for example, in essay and essay test assessment (Hamp-Lyons, 1995). The disadvantages to essay tests are that, because a time limit is imposed and because the students are writing in their second language, the sample may be very short. There are also concerns with the prompts provided; one concern that affects construct validity is the comparability of the various topic types used on the test (Stansfield & Ross, 1988). Some argue that restricted-time essay tests may produce writing that does not resemble the writing that the participant would produce under process-writing conditions (Caudery, 1990). Further, writing samples are time-consuming to score and the scoring is more subjective. Judging writing is precariously like grading personalities. One rater may consider the positive value of a certain characteristic to be so predominant in an essay as to overshadow a multitude of shortcomings, where another rater may disagree. Vaughan (1991) notes that while holistic assessment clearly works for some essays, in many cases holistic assessment does not work, and in these borderline cases, raters may well follow their own styles of judgment. Essay scoring also requires more expertise on the part of the scorers (Leki, 1991) although Kaczmarek (1977) says that both subjective and objective essay evaluation methods work well and each correlates highly with other valid ESL proficiency measures.

Research Question

Many IEPs use the TOEFL score to make admissions decisions (Leki, 1991) and for placement purposes. It is assumed that if an incoming ESL student achieves a certain score on the TOEFL, then she or he is ready for the content of a certain level in the IEP. And, if

1465

an ESL student is continuing in an IEP, then the TOEFL score is factored with the recommendations of the student's teacher(s) for placement. However, it is generally recognized that the TOEFL is not a valid measure of a student's position within a particular IEP (Leki, 1991). Instead, placement decisions should be matched with assessment instruments related to the IEP program curriculum. For this reason, the Center for English as a Second Language (CESL) sought measures that would be valid for placement decisions within its curriculum. Also, the TOEFL includes no sample of student writing and recommendations from teachers do not consistently include considerations of student writing.

Recently, Educational Testing Service (ETS) has required TOEFL testing sites to administer sample TWE essay tests. The TWE was developed based on survey data which revealed that faculty believed writing has a major role in the academic community and tests of writing should be based on what students are expected to produce (Kroll, 1991). Initially, the primary purpose of these tests is to train scorers so that TWE can become a part of TOEFL. (As of May 1990, ETS reported that the variance accounted for by the TOEFL section scores predicting the TWE score ranged from .80 to .84 (n=91,146) (DeMauro, 1993).

In addition, however, TWE sample essays can be used to make placement decisions. For example, at the time of data collected for this study, at the CESL, decisions regarding placement were made using TOEFL scores and/or teacher recommendations as information sources. The TOEFL was required when a student entered and exited the program. Course grades were used to determine progress through the program. A passing grade of 75% in a writing class, based on an average of in-class and out-of-class writing as well as a timed final exam, would allow a student to pass to the next writing class. (Split-level placement was then used.) Progress in other skills, such as grammar, reading, and speaking, was measured independently with the same score—75%—required for passing. However, using essays like the TWE to help determine placement is emerging as an option. At present, scores on the TWE are not used in placement decisions, and split-level placement has been abolished. The central question in this research study is whether to use writing samples as a means for determining placement and/or advancement in addition to relying on scores from objective tests and/or teacher recommendations.

1466

Method

Participants

The participants were 107 IEP students in CESL levels 1 (beginning) to 4 (advanced) enrolled in Summer 1993; 96 continuing students, originally placed by TOEFL score and advanced by passing course work; and 41 new students, placed by TOEFL.

Scores from 105 participants were used; 2 were dropped because of disagreement in raters' scores after 5 readings by 4 raters.

Test

The instrument used was the English Essay Pretest (EEP) from ETS, which was used as the pilot instrument for gathering reliable prompts for the initial versions of the TWE. So for the purposes of this study the terms EEP and TWE are interchangeable. The test was given in the third term (May-July) of the 1993 academic year. The test consists of a thirty-minute writing sample based on a single writing prompt. The graders used the 6-point criterion-referenced TWE scoring guide.

Raters

The raters were one applied linguistics professor, with 6 years of ESL teaching experience, and two IEP teachers, with 18 and 9 years of ESL teaching experience, respectively. All had had experience with teaching ESL writing and all were trained in the scoring of EEP/TWE essays. A fourth experienced rater was involved in the grading of the essays only during the last round.

Preliminary Calibration Set

The raters initially holistically test graded 7 final essays from CESL students written during the previous year and discussed their scores; differences were resolved based on the criterion-referenced scale. The actual grading was then done.

A grade correspondence as used in TWE ratings was implemented; that is, interrater agreement depended on two readers assigning the same whole number scores or +/- one whole number score to the essays (Kroll, 1991).

1467

Grading of the Essays

For the first reading, the essays were divided equally among the raters, who read them independently, using the TWE scoring guide and the sample essays for reference. A second round of independent readings was done, with scores from the first round being kept secret. Comparisons of the first and second readings yielded these results: Two raters had agreed on the scores for 94 of the essays (88%). The group of essays for which the 3 readers did not agree within +/- 1-point were given to the rater who had not read them (that is, if Raters #1 and #2 had not agreed on the score for essay #1, then that essay was given to Rater #3 to score). Of the 13 essays that required a third reading, all but 2 of them received a score from the third rater that was within +/- 1 point of one of the other two raters. The remaining 2 were sent to a fourth independent reader, but were finally removed from consideration in the results because no clear consensus could be reached regarding their scores.

Interrater Reliability

Because the essays were read in successive readings, an interrater reliability score for each round of readings was calculated. The first and second rounds combined produced an interrater reliability score of .879 and the third round, .846. The average interrater reliability score over the readings was .862.

Essay Score and CESL Level Correspondence

Correspondence between the EEP grading rankings and CESL levels was established by the two raters who are also instructors in the IEP (CESL). Although the instructors may have taught some of the participants and made recommendations for them, the evaluations of the EEP essays, which were known by number only, were rated on the TWE/EEP scale for which the raters established interrater reliability during the calibration session. Having finished grading according to the TWE/EEP scale, the two raters who are the teachers in the IEP reviewed the sample essays and decided that EEP scores 1 through 4 corresponded to placement into CESL levels 1 through 4, respectively; an EEP score of 5 would place a student into undergraduate study or into CESL level 5, for students preparing for graduate study, whichever was appropriate for the student. A student who scored 6 would be proficient for graduate study.

T468

Results

The correlation analysis for overall placement (independent variable: X) and EEP score (dependent variable: Y) is presented in Table 1. The correlation coefficient (r) is .58; thus, the placement of the student into a class accounts for 34% of the variance in the EEP score (r-squared), with a p-value of .0001, significant at the .05-level of alpha.

Table 1
Correlation of Placement (overall) and Writing Score (N=105)

Simple Statistics						
Variable	Mean	Std Dev	Min score	Max score	Corr	p
Overall placement	3.06	1.00	1	4		
Writing score	2.45	0.97	1	5		
Pearson correlation coefficient*					0.58	0.001

*H^0: Rho=0

The correlation analysis for placement by teacher recommendation (independent variable: X) and EEP score (dependent variable: Y) is presented in Table 2. The correlation coefficient (r) is .62; thus, the placement of the student into a class accounts for 38% of the variance in the EEP score (r-squared), with a p-value of .0001, significant at the .05-level of alpha.

Table 2
Correlation of Placement by Teacher Recommendations and Writing Score (N=63)

	Simple Statistics					
Variable	Mean	Std Dev	Min score	Max score	Corr	p
Placement by teacher recommendation	2.60	0.96	1	5		
Writing score	3.40	0.73	1	4		
Pearson correlation coefficient*					0.62	0.001

*H^0: Rho=0

The correlation analysis for TOEFL score (independent variable: X) and EEP score (dependent variable: Y) is presented in Table 3. The correlation coefficient (r) is .55; thus, the placement of the student into a class accounts for 30% of the variance in the EEP score (r-squared), with a p-value of .0002, significant at the .05-level of alpha.

Table 3
Correlation of Placement by TOEFL Score and Writing Score (N=41)

	Simple Statistics					
Variable	Mean	Std Dev	Min score	Max score	Corr	p
Placement by TOEFL score	423.36	63.28	308	543		
Writing score	2.19	0.95	1	1		
Pearson correlation coefficient*					0.55	0.001

*H^0: Rho=0

The correlations for the EEP with overall placement, with placement based on teacher recommendation, and with placement based on TOEFL score are statistically significant and moderately high. What is of additional interest is how the EEP would have fared as a predictor of class level placement. The researchers had predetermined the correspondence between the EEP score and the CESL level (see p. 111).

Discussion

The results of this study suggest that the TWE, or a similar essay exam, is a useful instrument for placement into and advancement within an IEP. In a good percentage of the cases, the EEP score was a more conservative measure than was either the TOEFL score or teacher recommendation when used to place or to advance students (Tables 4-6).

In the majority of TOEFL-score cases (17), students were placed by their essay scores at the same level as their placement by their TOEFL scores. The second largest group (11) placed one level below their placement by their TOEFL score on the basis of their essay scores. Other students were placed at one level higher than placement by TOEFL score based on their essay scores (7), 2 levels lower (6), and 2 levels higher (1).

However, the substantial number of cases in which students placed one level lower than TOEFL placement on the basis of their writing seems to reflect what is known from the experiences of classroom teachers: Producing written English is often a more difficult and complex skill than is choosing the correct form from several provided. The writing skill is often on the same level with overall English-language ability (Kaczmarek, 1977), but almost as often lags behind. This is a fairly common observation made by the researchers in this study.

In the case of placement based on teacher recommendation versus placement by an essay score, an even higher correlation was found, enhancing the reliability of the essay score as a predictor of global language ability. Again, though, placement of students according to essay scores was often at levels lower than placement according to coursework grades and teacher judgment. Placement according to essay test score was at the same level as teacher placement in 20 cases, and one level below teacher placement in 33 cases. Essay scores placed 2 students one level higher than did their teachers' recommendations; 7 students two levels lower than teacher recommendations, and one student three levels lower.

1471

Table 4
Comparison of Placement of Students Using Writing Score Only Versus Actual Placement (N=105)

	Writing Score				
Actual class placement	1	2	3	4	5
1	6	3	0	0	0
2	10	9	2	1	0
3	0	18	10	3	0
4	1	13	16	12	1

Table 5
Comparison of Placement of Students Using Writing Score Only Versus Actual Class Placement by TOEFL Score (N=42)

	Writing Score				
Actual class placement	1	2	3	4	5
1	6	3	0	0	0
2	4	5	2	1	0
3	0	4	4	2	0
4	0	6	3	2	0

1472

Table 6
Comparison of Placement of Students Using Writing Score Only Versus Class Placement by Teacher Recommendation (N=63)

	Writing score				
Actual class placement	1	2	3	4	5
1	0	0	0	0	0
2	6	4	0	0	0
3	0	14	6	1	0
4	1	7	13	10	1

If the phenomenon of a lower proficiency in writing is taken as a normal, consistent, reliable feature, as it appears to be from these results, then the discrepancy in TOEFL placements and placements by essay scores can be expected. It may be that teachers and administrators will recommend this lower placement. Or, only the essays that fall below the standard one level lower than TOEFL placement or above TOEFL placement level will be taken into account. In either case, the essays will be given serious consideration as reliable indicators and concrete evidence of the student's ability to use the language.

The issue of the two essays that were excluded from the study because the three raters could not reach agreement on scoring them deserves special consideration; namely, what was characteristic of these essays that rendered them "good" in the eyes of one rater and "not good at all" in the eyes of another? The researchers took a closer look at these two essays to understand the process of scoring these two essays in particular. The first essay received the following scores from the raters: Rater #1: 3; Rater #2: 2; Rater #3: 4. These scores held for a second round of reading both within a month of the original readings and a year afterwards.

Comments from Rater #3, who gave the essay the highest score, were that the essay was well-organized and content was treated logically; the essay was written with good grammar; and there were "no terrible handwriting or spelling obstacles." Rater #1, who scored in the middle, commented that there was only a weak thesis statement which affected the essay's organization; sentence structure and phrasing were wordy and awkward; one of the main points in the essay did

1473

not have enough detailed support; and word choice was weak in places. Rater #2, who gave the essay the lowest score, said that word choice was simplistic and, at times, incorrect; many run-on sentences appeared; there was no evidence of knowledge of conventional formatting; and none of the points was developed well.

In the case of the other essay, the following scores were given: Rater #1: 2; Rater #2: 1; and Rater #3: 3. The highest score was given by Rater #3, who commented that the essay was thoughtful, fragmented and, well-organized and it contained significant grammar problems. Rater #1 said that the essay contained no clear thesis, had a lack of transitions and a lack of detail. This rater also noted the sentence fragments, errors in tense, and awkward sentence structure. Rater #2 commented that the writer tackled two big issues in an essay of essentially only two paragraphs and, thus, was underdeveloped. In addition, this rater noted serious grammar errors, including run-on sentences, fragments, shifts in verb tenses, and virtually no use of the article system.

From the results of this study, the researchers concur that, because essay exams require a productive use of language which integrates a wide spectrum of linguistic skills, such exams provide a more robust, albeit more complex, measure of English ability than does a test consisting largely of recognition and response to structures already provided, as is the case with the TOEFL. Writing an essay, like carrying on a conversation, requires the student to recall and integrate vocabulary and structure in order to create meaning. Thus, such a sample of language gives more useful information about the individual's skill in using the language than does an objective test of discrete items.

We agree that the rating process in writing assessment needs further research regarding scales and forms of evaluation and, indeed, more attention paid to the shared values of raters regarding the language and content relationship in writing (Connor-Linton, 1995; Mohan & Low, 1995). At the same time, we would argue that essay scores can provide an acceptable measure for student placement when they are used in combination with TOEFL scores (and/or another measure, perhaps an oral test such as the Foreign Service Interview) and that they add a dimension to assessment that is not available with the use of an objective measure such as the TOEFL alone.

References

Caudery, T. (1990). The validity of timed essay tests in the assessment of writing skills. *ELT Journal, 44*(2), 122-130.

1474

Connor-Linton, J. (1995). Looking behind the curtain: What do L2 composition ratings really mean? *TESOL Quarterly, 29*(4), 762-765.

DeMauro, G. (1992). Examination of the relationships among TSE, TWE, and TOEFL scores. *Language Testing, 9*(2), 149-161.

Diederich, P. (1974, Spring). *Measuring growth in English.* Paper presented at the meeting of the National Council of Teachers of English, Urbana, IL.

Educational Testing Service. (1992). *TOEFL: Test of Written English Guide.* Princeton, NJ: No Author.

Hamp-Lyons, L. (1995). Rating nonnative writing: The trouble with holistic scoring. *TESOL Quarterly, 29*(4), 759-762.

Kaczmarek, C. M. (1977). Scoring and rating essay tasks. In B. W. Robinette (Ed.), *Papers in ESL: Selected conferences* (pp. 151-159). Washington, DC: NAFSA.

Kroll, B. (1991). Understanding TOEFL's Test of Written English. *RELC Journal, 22*(1), 21-33.

Leki, I. (1991). A new approach to advanced ESL placement testing. *WPA: Writing Program Administration, 14*(3), 53-68.

Mohan, B., & Low, M. (1995). Collaborative teacher assessment of ESL writers: conceptual and practical issues. *TESOL Journal, 5*(1), 28-31.

Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly, 17(4)*, 651-71.

Perkins, K., & Pharis, K. (1977). TOEFL scores in relation to standardized reading tests. In B. W. Robinette (Ed.), *Papers in ESL: Selected conferences* (pp. 142-61). Washington, DC: NAFSA.

Stansfield, C. W., & Ross, J. (1988). A long-term research agenda for the Test of Written English. *Language Testing, 5*(2), 160-86.

Vaughan, C. (1991). Holistic assessment: What goes on in the rater's mind? In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 111-125). Norwood, NJ: Ablex Publishing.

White, E. M. (1993). Assessing higher-order thinking and communication skills in college graduates through writing. *Journal of General Education, 42*(2), 105-122.

1475

Authors

LYNNE DAVIS, Instructor, Center for English as a Second Language, Department of Linguistics, Southern Illinois University, Carbondale, IL 62901-4518. Email: ldavis@siu.edu. Specializations: writing instruction, spelling instruction, grammar, ellipsis in spoken English.

RUTH JOHNSON, Assistant Professor, Department of Linguistics, Southern Illinois University, Carbondale, IL 62901-4517. Email: rjohnson@siu.edu Specializations: intercultural communications, pronunciation, alternative assessment.

FLOYD OLIVE, Instructor, Center for English as a Second Language, Department of Linguistics, Southern Illinois University, Carbondale, IL 62901-4518. Email: folive@siu.edu. Specializations: L2 writing, CALL, alternative assessment.

1476

News and Views

Applied Language Learning
1998, Vol. 9 Nos. 1 & 2, pp. 121-145

Interpreter in Action
Interview with Lieutenant Colonel Richard Francona, U.S. Air Force (Retired)

Lidia Woytak
Defense Language Institute Foreign Language Center

The star-studded Pebble Beach golf tournament was a few weeks away. It was still a sunny afternoon: Students in camouflage uniforms were rushing upstairs to the top floor auditorium. Teachers were heading for seats in the back row. Sounds of Arabic speech filled the room. Deep into the sky, one could faintly trace the sandy expanses of the desert. In some distance, a shadowy figure in an Arab garb appeared on a horse. Was it Lawrence of Arabia? Was he descending to hear his name evoked in this room?

Thanks to Mr. Bahgat Malek, a department chair of the Middle East School 2, who organized the visit of the retired Air Force Lieutenant Colonel Rick Francona, these students as well as their teachers now had the opportunity to enrich their knowledge about the job of an interpreter participating in complex multinational operations taking place in the Middle East. Francona was no stranger to DLIFLC. In 1974, he was a student of the Basic Arabic Program at the Institute and in 1978 he returned to work here as a Military Language Instructor. To this day, he considers the DLIFLC Basic Language Program superior to all other programs.

Introduced in both Arabic and English, Rick Francona emerged amidst a burst of applause to transpose the students and teachers from the safe environment of homework, tests, and exercises into the world of military operations in the Middle East. The timing for inviting a former military interpreter to DLIFLC was perfect. Today, security of the world is challenged by regional conflicts which transform regions into states and states into regions. The number of international conflicts over land, oil, and water is growing: Most of them are anticipated to continue. The United States, a United Nations (UN) member, will inevitably remain globally involved in the resolution of such conflicts.

Throughout his career, Colonel Francona was actively involved in international affairs. It began in 1976 when he participated in the

1477

evacuation of the US Embassy in Beirut, Lebanon. Subsequently, he became an advisor to the Royal Jordanian Air Force in Amman, Jordan. Following his tour at Headquarters, U.S. Air Forces-Europe, he became a liaison officer to the Iraqi armed forces Directorate of Military Intelligence in Bagdad. Following that country's invasion of Kuwait, Colonel Francona was sent to the Gulf as the personal interpreter and advisor on Iraq to Commander-in-Chief of the U.S. Central Command, General H. Norman Schwarzkopf. In March 1991, he served as an interpreter in the Safwan cease-fire talks in Iraq.

Following a brief review of historical events in the Middle East in the recent past, Francona focused on major economic, political , and military factors defining the area. He explained that access to oil at reasonable prices is a major undercurrent of international interests. The images of Operations Desert Shield and Desert Storm were circling in listeners' minds as this regional expert drew a comprehensive picture of the events in the Gulf.

Francona's experience in the Gulf entailed a dramatic twist: From serving as an Iraqi ally during the war with Iran to becoming an opponent following its invasion of Kuwait. He described the drama of dealing with the same individuals, first as friends and then as enemies in his soon-to-be published book titled *Ally to Adversary: An Eyewitness Account of Iraq's Fall from Grace*. In the book, he guides the reader from the final Iraqi offensives in the eight-year ling Iran-Iraq war to the Iraqi defeat at the hands of the U.S. lead coalition in 1991. He also describes his experiences in dealing with the Saudis and other members of the Coalition.

During the lecture Colonel Francona frequently resorted to humor. The students burst into laughter hearing that he was not sure how to react to General Schwarzkopf's greeting, "You're the Air Force." No one could surpass his humorous description of the pompous arrival of the Iraqi generals at Safwan for the cease-fire talks. He was incredulous when the deputy chief of staff of the Iraqi Army asked him, "Which one is General Schwarzkopf?"

In many ways, Colonel Francona personified the Warrior of the 21st Century as depicted in *Joint Vision 2010:* dependable, loyal, and ready to serve. Francona did not need to describe his military readiness. He just said, "In the middle of the night I got a call telling me to be on the plane to Iraq the next day. I was." Colonel Francona advised the DLIFLC students to be prepared for an important mission ahead of time. He said, "When it comes, you have to be ready."

According to Francona, a solid language foundation and background knowledge of the area are two major elements defining a

1478

successful career of a military linguist. During the lecture he encouraged students on several occasions to learn about the area on their own as much as they can. He added that as a student at the Institute, he found the *Area Studies Handbooks* for the Middle East very useful.

Colonel Francona also pointed out that teamwork is essential. He said that many times he was in situations in which his team members could not do the job without helping each other. Whenever the interpreters on his team needed an expert in Saudi dialect, they would turn to a colleague, for advice. Francona's personal account gave the students an overview of everyday lows and highs of the main duties of an interpreter.

Francona's presentation gave also an opportunity for the faculty to reflect upon the scope of duties of their graduates as well as their long journey from the first day in a language department to the festive graduation. Does this journey still lead to fulfilling today's needs? Do portions of the syllabus call for a readjustment? What comes first: the chicken or the egg, Modern Standard Arabic (MSA) or Iraqi?

The lecture of Colonel Francona made the listeners aware of the difficulties facing the Arabic students in the Institute's two Middle East Schools. Namely, they have to be able to select the language suitable to the text. They are tested on proficiency levels designed for one language—MSA. Yet they have to alternate between the MSA and one of several dialects to respond to a particular task. For example, Arabs would discuss current affairs in MSA, but they would conduct small talk in the Egyptian or Gulf dialects.

Notably, Richard Francona pointed out that small talk that many native speakers consider easy, is hard to master for the learner of Arabic as a second language. He stated that the unpredictability of small talk topic makes preparation for it difficult because anything may come up in an informal setting. Francona also pointed out that sometimes you cannot understand the other person without understanding some background information. While describing his interactions with the Saudis, he pointed out some speaker's tendencies to bypass what they believe to be common knowledge and focus only on new information. However, Francona amplified, what is obvious to a native speaker may not be obvious to a foreign speaker and thus may cause a breakdown in communication.

During the interview, Francona stated that teaching conversation on a one-on-one basis in his experience has been the most effective. He also recommended that students be given an opportunity to watch encounters of native speakers which could be acted out by two teachers. Such encounters, although expensive in terms of time to create, would

479

provide not only verbal but also kinetic models for the students to imitate. Videotaping then replaying such encounters would cut down on the expense. The basic encounters videotaped in the mid 1980s for the DLIFLC's Russian Program turned out to be successful. Natalia Goroshko and Leonid Slutsky presented their ideas on incorporating teacher encounters into curriculum in the article titled "Four-Handed Teaching." [2]

The interview with Francona makes the reader aware of complexities of skills required of an interpreter. Francona stated that while interpreting, he did not have time to think about the language. On the contrary, any time he attempted to think about the language, he found himself two sentences behind. In several instances, he alluded to the need of training in interpretation. A similar view was expressed by M. Kuwahata in "Sink or Swim: Five Basic Strokes to Consecutive Interpretation" presented during the Conference on Tradition and Innovation in Translation and Interpretation (MIIS, 1999).[1] In it she compared training in interpretation to training in swimming. Just how do we learn to swim? Is it more effective to be thrown into a deep pool so you have to somehow find your own strategies, or is it better to be taken step by step and led gradually through the various skills?, she asked. As in swimming, military linguists need training in interpreting. Daniel Gile in his book *Basic Concepts and Models for Interpreter and Translator Training* (John Benjamin, 1995) recommends activities aiming at development of comprehension abilities and verbal fluency. A completion of a rigorous six-months course in medical or court interpreting leads to a certificate. Is there a need of such a certificate for the military personnel?

During the interview, Francona stated that interpretation is an important part of military operations. Interpretation, recognized as a skill at the 1919 Paris Peace Conference and focused upon again during the famous 1949 multilingual Nuremberg trials, has also become an important part of the global language industry. In the 80's, former Congressmen Leon Panetta and Paul Simon wrote bills (H.R. 2608 and H.R. 3029) that brought interpretation and translation into the forefront of political agenda. Currently, the American Society for Testing and Materials Committee, F15.34, is reviewing national standards for language interpreting in the United States. Research on community interpretation in medical and judicial establishments indicates that not only interpreters but also personnel they work for should be aware of the basics of interpretation.[3] Perhaps the military could also apply some findings from this research.

Lieutenant Colonel Richard Francona has shown these student-

1489

soldiers, sailors, marines, and airmen that a military linguist adopts "language for a lifetime." Although he retired from active service in 1998, he did not end his career with a retirement luncheon. He is just now beginning the most creative period of his post-service career: publishing memoirs, giving lectures, and, most important, teaching young service members and encouraging them to "be all they can be," to "aim high," and to make it an adventure and not just a job.

*
* *

Welcome, Lt. Col. Rick Francona and Major Emily Francona. It gives me great pleasure to talk to both of you. Colonel Francona, tell us about the content of your new book titled Ally to Adversary: An Eyewitness Account of Iraq's Fall from Grace *and why did you decide to write it?*

The book deals with my service as the CENTCOM interpreter in Operations Desert Shield and Desert Storm. It tells the story of my first experience in Baghdad in 1988 through a few months after the end of Desert Storm. It's a story of my observations of how Iraq changed from being a U.S. ally to one of our primary enemies in the latter half of this decade.

I wrote the book to present the history of Iraq as a personal story rather than a scholarly journal or as a one-sided correspondent account. There is always a temptation to pitimise your enemy. I wanted to tell the story, how we have come from where we were to where we are now. I wanted the book to be an easily readable, personal story–not a textbook.

To what type of readers is your book directed?

It is directed to the lay, general audience. I would imagine most people interested in it would be military people, linguists, and faculty and students from the schools here. It has very little jargon and few technical terms so the general readership can follow it. It is meant to be an entertaining, yet informative story.

Mr. Malek told me that the Iraqis did not want to deal with you initially because your language was so excellent. They thought you were a native.

They thought I might have been a Lebaneese-American. The Iraqis were

1481

surprised that I would prefer to conduct business in Arabic. Even though one can speak Arabic, they would always provide an interpreter and begin to conduct the conversations in English. I felt more comfortable talking with them in Arabic than in English: It turned out to be about half and half. I preferred speaking in Arabic so that there would be no misunderstanding of what we were talking about.

I was concerned about their understanding of the English; and they were concerned about my understanding of the Arabic. Many times we would do it in both languages, trying to come to terms in both languages. These conversations were dealing with ongoing U.S. and Iraqi military operations. It was very important that they were done correctly.

What was your official function in Operations Desert Shield and Storm?

I was the U.S. Central Command's interpreter. We had several people, but none of us was a certified interpreter. I was responsible for getting things done. I couldn't be available 24 hours a day. I would be gone for four or five days talking to deserters or Bedouins. In the meanwhile, four or five others handled daily chores that needed to be done back in Riyahd.

I coordinated where people went, even officers senior to me. We had native speakers available in the U.S. forces, very capable officers and warrant officers. I made sure that every general officer that needed an interpreter had one available. I also made sure we had enough interpreters to rotate, because interpreting can be very nerve-wracking and intensive. We could only do it for so long before we would start loosing the edge: no more than 45 minutes without having another interpreter take it up. We would sit in the room when not interpreting, so we were not brought in cold. I don't know every Arabic word. If I didn't know a word and presumed it to be important, I would write it down and pass it behind me. We would all check each other's interpretations to make sure everything was covered and also to back each other up. It was a team effort.

How long can an interpreter perform fully in simultaneous interpretation?

After about half an hour, you start getting really, really tired because you are listening and talking almost at the same time. You never have time to formulate what you want to say. The moment you stop to think about the meaning, you are already two sentences behind.

1482

An interpreter's credibility is an important asset. How did you establish your credibility as an interpreter?

I already had some credibility with the Iraqis. I was known well enough to the CENTCOM staff that they asked me to come over. Then they asked me who else we needed. At that time, CENTCOM could pull anybody they wanted. I gave them four names and said, "You need these people; they are the best I know." They all performed well.

How did I establish my credibility? I established it initially with the Saudis. When we were not interpreting, we did a lot of correspondence with them. I proved to them in my daily work that I was capable of performing that task. They came to rely upon me. Thus, when there was a meeting or correspondence, they would say, "You have a good interpreter here, we don't need ours." That was how I became known. The senior Saudi staff asked for me to brief them daily in Arabic. Subsequently, I was also called on to brief the King and Minister of Defense and the King.

Could you describe the exchange of communications between the Arabic and English-speaking sides in the Gulf region?

There were very few formal negotiations or consultations with the Saudis, but there was a lot of informal verbal communication that needed interpreting: mostly informal briefings, conveyance of information, and routine exchange between the forces. Everything formal was sent via a memo or a letter. They would write to us in Arabic and we would reply in English so we were always generating our own language. The written communications we would do were fairly routine in content and vocabulary. I would do voice interpretation for the general during an informal meeting or a courtesy call. For example, the Egyptian commander would come down from his unit to visit the headquarters to pay a courtesy call on both General Khalid and General Schwarzkopf.

Could you describe the initial negotiations with the Iraqis?

The verbal exchanges stayed focused because while the translations were occurring, the principal was already thinking of the next utterance he was going to say. There was always a lag time. There was nothing said that should not have been said. Everything was thought out in advance. As we knew both the talking points and the agenda, we had a general idea of what General Schwarzkopf was going to say.

DTIC 483

We recorded on cassette tapes formal negotiations with the Iraqis. We had four tape recorders to make sure that everything was recorded; a copy of the tapes would go to the Iraqis, and a copy of the tapes to the Coalition. There were three of us sitting at the end of the rectangular table. General Schwarzkopf and General Khalid were to my left, the Iraqi officers to my right, and the Iraqi interpreter across from me. It went in a circle. The Iraqis officers looked at each other: General Schwarzkopf and General Khalid would talk to each other. We would speak in Schwarzkopf's direction. Most of the time the opposing generals maintained eye contact with each other trying to gauge each other's reactions as soldiers do. They had time to collect their thoughts and make sure they were saying exactly what they wanted to say. It worked out well and went quickly.

Were you involved during the Second Safwan Talks?

We maintained contacts with the Iraqis almost every day during the Second Safwan Talks. I would go along for the headquarters element if a general officer was to speak. I was not the only one: There were hundreds of interpreters out there. Every American unit had its interpreters. They all did really great jobs.

One time we were interpreting a long, detailed conversation about the repatriation of 80,000 prisoners of war from Saudi Arabia back to Iraq, but this time between an Iraqi and a Saudi, both speaking Arabic. Since the U.S. officers did not understand Arabic, we were interpreting both sides of the exchange simultaneously.

We were "ganging up," as we called it. I was on one side, my colleague Vernie, on the other. I would interpret the Iraqis and he would do the Saudis to give the principal the idea of a two-way conversation. Later on, they were getting down into the specifics of how many buses per day, per hour, and what border crossings. This level of detail wasn't necessary for the U.S. general officer. Since he had a lot of other things on his mind, I gisted the conversation for him. I would say "The Iraqis are saying they need to put *x* number of buses out and the Saudis are replying that it was not enough." At the time it was sufficient to summarize the conversation for him because, later on, these conversations were going to be transcribed and typed.

Did you experience any difficulties in handling formal communications?

The more formal they got, the easier they were. Briefings in which we

1484

conveyed significant information in a short amount of time were probably the easiest because they contained prearranged information. We had the briefing format on plastic boards and we would, in grease pencil, update the relevant information by assigning new values to the data. We gave the briefing two or three times a day in Arabic to different audiences. I could probably still give you that briefing today because they were so formalized.

Did you enjoy giving the briefings?

Yes, I did. Dealing in the military environment was probably the easiest because I was familiar with it. When I was talking to someone in a uniform, I could depend on visual clues. I could see what rank he was, so I would know what education he had, and at what level of responsibility he held. The badges would tell me what experiences he had and in what branch of service. I could adjust my speech appropriately.

Did the Saudi military use a standard memo form?

The memos from a Saudi commander or his staff were on a form. The form had its basic elements which were always the same: The date would be on the right, the serial number would be on the left, who sent it, subject, paragraph one, two, three, four, and then the signature block. It was just like a U.S. memo. The first paragraph would contain a formal greeting. However, once they got to the subject, it could be anything.

Would they get into the subject in the beginning or in the middle of the letter?

These were military people, trained in U.S. schools. They followed the U.S. standards in which the subject was followed by a statement of need. One of the most important things we translated into Arabic was the war plan. We took a lot of time to translate it because it was very detailed, specific, and critical. The Saudis made their comments on the war plan in Arabic. Afterwards we responded to it.

You encountered such speech events as negotiation, events update, discussion, address, letter, and small talk. What type of speech event was the most difficult to interpret?

Small talk. I dealt with Saudi small talk in the Gulf during Desert Shield and Desert Storm. The only small talk I would have had with any Iraqis

485

would have been with some shop owners when I was able to get out and move around. It was difficult as the Iraqis would neither drop their surveillance nor would they let me move around by myself—although at times I would sneak away. I had a lot of small talk with my colleague about his family, his wife and kids, and similar topics. I found talking in a dialect with the children and the wife who spent all day in the house challenging and entertaining, too. My colleague made gestures to help me understand.

Small talk was the most challenging linguistically because it could be about anything. During what I would call "down time," I was often called upon to interpret, especially during Desert Shield when we were at the headquarters waiting for the forces to arrive. They would say "Please explain to the carpenters that we need them to cut a 2x4-foot panel and to order pieces of cork." I had not often come across trade expressions. Responding to the call, I learned a lot of trade expressions that I would not have learned otherwise.

In every small talk you hear certain phrases over and over again. The more numerous and precise they are, the less capacity you need to handle them. They can free productive capacity for creative interpreting. Do you think we should train students to supply instant phrase equivalents?

Yes, rather than having to translate the phrase, they would know what the phrase is. There are a lot of those in Arabic, although, they differ from country to country. The phrases are key.

Also, abbreviations are very important and they are used more and more.

Sometimes you have to paraphrase because you don't have an equivalent.

During my lecture at DLIFLC, I told the students an anecdote about paraphrasing. After a long negotiation regarding U.S. military women driving in the Kingdom, the Saudis announced their king's decree, "American military women driving military vehicles in military uniforms are *not women*." Should we know a literal translation of that? Yes, of course. Should we use the literal translation? Obviously not. We reformulated this decree into a less offensive statement, "American military women may drive vehicles while in uniform." Same information, but not a literal translation.

1486

Do Arabic speakers use a lot of idiomatic expressions?

Absolutely. They use many idioms in conversations, but not in formal communications. Frequently you understand the words, but the meaning makes no sense to you.

Let me share an anecdote with your readers. Whenever visiting congressmen and generals would want to buy a souvenir of Saudi Arabia, I would always be tasked to take them downtown. We would go to different merchants—they all were capable at taking money in any language. I would always tell the visitors, "I will interpret for you, but I will not negotiate for you. When this is done, you all have to agree on the price. Haggling is your job, I will merely mention the numbers." They said, "Okay, this sounds like fun." It was entertaining for them, and we would go downtown and the merchants would say, "I'm going to give you the best price." I would translate it. Then the buyer would make an offer and the merchant would reply saying, "Now listen, we're just talking bread here, not butter." And I would translate that, "He's talking bread, not butter." and I would always get this "What does that mean?" look. I would say, "I'm just telling you what he said." Then I would explain, "He's only making the bare minimum here. At that price, he would not make enough to afford butter." For a lot of these I learned standard answers, but in official business, I rarely encountered them.

What kind of expressions posed a problem for the U.S. military interpreters?

In general, technical vocabulary, military concepts, concrete terms were easy to interpret or to translate. We had problems, however, with phrases and expressions that had a cultural meaning. When the Saudi general officers were describing the morale of the forces during the war-planning discussions, they would refer to Arab history or Islamic concepts, rather than the military ones. It got confusing for us. These concepts were familiar to them because they learned them through their socialization, in school or at home. Sometimes we would get lost because we knew the words, but did not know the meaning. Other times, they would refer to another person saying, "Remember the story of *so and so*." We had never heard the story of *so and so*.

To them the story had carried a certain meaning.

Whenever we talked this way in English, it was also confusing to them. We had to decide that either this segment was important enough to stop

1487

and ask them what they meant, or ask them if they would send it to us in writing later so that we could figure out what they meant. In their formal writing, they would not use historic or religious terms. The translation that came to us was pretty standard and stuck to the business at hand. In their free flowing conversation, however, we frequently found unfamiliar shortcuts in references to experiences that they shared.

They relied on their background.

Just as we do. We would make little quips to each other; punch lines from common jokes, stories, or English proverbs. They would do the same thing. They would recite a phrase out of the Koran which they thought conveyed the meaning. Not being conversant in the Koran, we didn't get it.

Did they sometimes use the Saudi dialect to talk in private?

No, not the Saudis. I don't recall any instances of using language as a tool to hide something from us. I know they could have. Most of the time, they were in a communicative mode. In the Command Post, we had enough understanding that if the two Saudis beside me wanted to keep their conversations confidential, they would get up and leave rather than switch to a dialect. Although many times, if they were deep into some dialect, they could have sat right there and had the conversation.

What people do with words and how they act differ from one group to another. Have you observed that the same stimulus can trigger different reactions from Americans as opposed to Middle Eastern people?

At the Institute, students learn from native Arabic instructors what is acceptable in the Middle East: what subjects are not joked about and what phrases are not accepted in common speech. Students get training here on politeness and some follow-on culture training. However, the principals do not always get this training. Sometimes they opened up a meeting with a culturally inappropriate phrase.

Frequently, officers would ask improper personal questions of the Saudi officers about their families, about their wives and children. Wives are a subject best left undiscussed; children are okay. You should know someone before you start asking personal questions, because the Saudis are private people and they do not want to discuss these subjects. Although the familiar term of address in the Arab world is to call someone *the father of* and then give the son's name, but one should not start

1488

off with such familiarity. A lot of times the Americans would say in a friendly manner "Hi! How are you? Do you have a wife and kids?" or the really personal, and offensive, "How's your wife?" I have been trained, so I do not ask. If you do ask, the Saudis would probably reply, "Oh yes, I have a son. His name is Mahmud." Then one might think, "I'm going to be really friendly. I'm going to call this guy *Abu Mahmud*." One should not do that right off the top. When an interpreter is introducing a U.S. general to a Saudi general, and the former starts with this line of questioning, "How is your wife?" and "How are your children?" I would say "How is your family?" The Saudis know us: They know how we are, so most would not take offense anyway, but I would try to diffuse it by avoiding offensive exchanges.

Also, the subject of American politics in the Middle East is very sensitive. Most Arabs regard our support of Israel as problematic. A discussion of this subject has to be done diplomatically. Frequently we would use words and expressions that are inflammatory to an Arab. Whenever the subject came up, I would soften the translation to avoid a problem. Nowadays, our senior officers are gaining better awareness of the deep-rooted sensitivities to area politics.

Could you give an example of an inflammatory reference in politics?

A reference to Israel's right to exist citing U.N. resolutions is inflammatory. It brings up a whole range of retorts because the Arabs are well versed in all of the U.N. resolutions that the Israelis have ignored.

When you hear an utterance, you are focusing on the message. Besides the language you hear, what else do you take into consideration to formulate your message?

You look at the body language and the tone of voice. The Iraqis are like any other people; and personally I find them witty. They displayed a sense of humor even in the situation they were in. They often made quips that I found amusing. If translated, they would make no sense to the principal without an explanation that he was trying to be witty, or sarcastic, or derisive of the United States. I kept a notepad to make notes for the general. Once in a while one of the Iraqi generals would go off on a rhetorical comment about "We don't understand why the United States forces are still occupying part of our country. After all this was about Kuwait and we have left Kuwait." I would gist that for the principle. Then I would lean over on the table and write "BS" or "ranting" or "party line." So the general could look down and know this was

1489

rhetoric; not the substance of the talk. We always had to listen to this two-minute blast of rhetoric prior to getting down to business. The Iraqis were doing it because they were told to do it, not because they believed it. For the most part, I detected no disrespect on either side: It was just something we went through. They all were initially unfriendly, but after we set up the system they, for the most part, were responding respectfully and professionally. Except one time, one Iraqi officer was very arrogant.

Why was he arrogant?

I'm not sure. We met him at Safwan. Colonel Dunn had been dealing with this general on a weekly basis. When we saw he was coming in to represent the Iraqi side, the colonel called me aside and said "This guy is always hard to deal with." I asked "Who is he?" He replied "He is the Regional Commander and takes it personally that we are occupying his country." I commented "I guess I can understand that." I made some mental notes about the Iraqi and then briefed my general. I explained that the Regional Commander was known to be hard to deal with. The general's response was "I don't care. I'm up here to get some business done and we're going to get it done." It worked out.

Language utterances manifest themselves through body language. For example, smiles of Japanese students during a lecture indicate that they are merely attentive, not necessarily agreeing. Did you occasionally get a message expressed through physical movement?

Not really, except for one occasion. We had a very long meeting with the Iraqis at the Second Safwan in which I knew the Iraqi interpreter. He was a friend of mine: In 1988 I had worked with him for months. I had been to his house, I knew his family, we had gone out to social outings together (I'm sure they were all sponsored by the Iraqi government). I knew his gestures and mannerisms. So when he was translating some rhetoric about the Iranian Revolutionary Guard Corps infiltrating Iraq, true or not, I knew this was propaganda, because he was tilting his head in a certain way. I thought, "Oh, there he goes again. He's off with the Ba'ath Party again." I could tell by his facial expressions that his heart was not in it. I looked at him, I'm sure he knew my looks and gestures as well, and gave him the "Come on. Can we speed this along?" look.

I had more awareness of Iraqi body language than Saudi. I had to learn the Saudi gestures and nods when I got there. We were there in a cooperative atmosphere: If I did not understand something, I could

1490

easily learn it.

From the interpreter's point of view, what is the difference between the dynamics of a group versus a one-on-one talk?

Most of the time we were one-on-one. The only group settings were formal meetings or the talks with the Iraqis. I didn't really notice a change, because I was there to serve the general or senior officer present. I focused on what he wanted and what he needed; the other people in the room were his staff. Usually, we had interpreters for the U.S. officers if they needed them. At the Second Safwan Talks, however, I remember that the other U.S. officers at the table had no interpreters available and there was nothing we could do about it. Colonel Dunn and I were there to interpret for General Johnson: We gave him an ongoing account of the situation.

Would the opposing parties argue a particular meaning of a term during the negotiations?

Yes, in one instance General Schwarzkopf proposed that we establish a line on a map from which both sides would pull back by a kilometer to make a two-kilometer buffer zone. He drew a line on a map and the Iraqi interpreter used the term that meant a *political separation* rather than a *military cease-fire line*. The Iraqi general took great offense at what he thought was the United States drawing political borders inside his country. He argued with Schwarzkopf. He said, "No. I do not have the political authority to alter the political boundaries and you should not even be in my country. We can't be talking about solutions and demarcation lines." Schwarzkopf replied, "No. I'm not talking about that. I'm just talking about a line from which we can pull our soldiers back so they would not be killing each other." This exchange went back and forth for several minutes, until we finally said, "Okay, it is a line without political consequences. It is merely a line from which the military forces will withdraw." The Iraqis were concerned that the former meaning would have long-range political ramifications.

Describe the conversational style of the two opposing officers. What similarities and what differences have you observed?

The United States and the Saudi officers spoke to the Iraqis clearly, slowly, and simply. They avoided complicated terminology. General Khalid spoke to the Iraqis very distinctly in Modern Standard Arabic.

1491

General Schwarzkopf spoke in clear, precise English.

The Iraqis, on the other hand, spoke to us in a conversational tone. In contrast to our formal, almost stilted style, the Iraqis' style was informal.

Could you describe the speech patterns of the enlisted personnel?

I found the Saudi officers easier to understand than the enlisted because of their education level. The more education you have in the Middle East, the closer your speech comes to Modern Standard Arabic. Many of the younger troops or the Bedouin that we talked to had almost no formal schooling. Iraq has a high literacy rate. Say what you will about the Saddam Hussein government, but the Ba'ath Party has raised the educational standards in both Syria and Iraq. The Iraqis are quite literate. For the most part, everybody was understandable, but occasionally you would get some less educated people. They were hard to understand, because they only knew the Iraqi dialect that they learned at home.

Saddam Hussein is getting better and better at tailoring his style to the circumstances. For example, in his speech to the nation on the Gulf crisis, he told an elaborate story from Muslim mythology. On the other hand, his conversations with his generals consisted of simple and short sentences. How important is it for an interpreter to be cognizant of different registers?

I have always had problems listening to Saddam Hussein's speech because it is very difficult. But listening to Yasir 'Araphat's speech is much easier, because he speaks in a simple style. He does not use the mythological or religious metaphors. Saddam is more of an orator who is trying to arouse Iraqi nationalism. His style is very similar to Fidel Castro's and reminiscent of Gamal Abdul Nasser who could fire up people through his rhetoric. How much information is Saddam Hussein conveying in these speeches? The best thing to do with them is to allow native Arabs who understand all the metaphors to translate them, and get the content from a transcript. When I read the transcript in English, I find that I have missed nothing because he said nothing of importance. On the other hand, when he is passing out information and decrees, he is to the point, crisp, and easy to understand.

Have you ever met Saddam Hussein?

No. The closest I got, I was in the same bar as one of his sons. We were

1492

about eight feet away and judging from the physical security that was around him and the toughness of the cadre that was with him, I decided that was the closest that I wanted to be.

You mentioned in your lecture to the DLIFLC students that there is an influence of Soviet culture on the Iraqi system. In which areas do you see similarities?

Style of government, some of the government functions, some of the intelligence services, and the security services are similar to the former Soviet system.

On the social side, the Iraqi culture in large is not that affected by the Soviets or the Russians. The Iraqi military is not organized as a carbon copy, or a smaller version of the Russian armed forces. The Iraqis have taken what they liked from the British, because the British mandate is part of their heritage, and what they got from the Russians and they have incorporated both into their own unique model. There is a mixture: They use a lot of Russian tactics and planning, but the organization of the forces is western. Iraqis have done a good job over the years of taking the best from various cultures.

Could you focus on the difference between content fidelity and linguistic fidelity? Which one do you consider more important?

Content. When I do any kind of translation or any kind of interpreting, I always strive to relay the meaning, not the words. I would always tell the general what the speaker meant to say, not the words he actually said.

There are idioms and phrases that do not translate well. For example, the *mother of all battles*, a literal translation from Arabic has became a popular phrase in English. It should be translated *greatest of all battles* or *battle of battles*; that's what it really means. The *mother of all battles* is catchy; now it is used in all sorts of contexts.

Some sentences render themselves to more than one interpretation. What made you pick one and not the other?

Context, and here we go back to what we were talking about earlier regarding expectations to discuss certain topics. I would always go with what I felt fit the context of the conversation.

How did you interpret ambiguous utterances?

1493

I didn't. I would ask for clarification, or if it was something unimportant, I probably would let it go. If it was in an environment I was controlling, say I had an Iraqi officer, a cooperative deserter, I would ask him questions. As long as he was giving me the information that I was seeking, and he would mutter something under his breath, for example, "Well, that's life in Baghdad," I would let it go, because it wasn't germane to what I was trying to accomplish at the time. But if it was something about the subject, for example, he would point at a map and smirk, "Then there are these guys over here." I would say, "What do you mean *these guys*?" He would respond, "Don't you know?" .... and he would use a certain term. If it was different, I would say, "What do you mean by that?" He would say, "Well that's what we call the intelligence, the security guys." I said, "Tell me about them." because I didn't know. I had to decide what I was going to pursue.

In other words, you were searching for specific information.

I only got involved when there was a prisoner or a defector, or a deserter who had unique information that was pertinent to our activities. In one instance, we were trying to find a CBS journalist that had been captured by the Iraqis, Bob Simon. One of the prisoner-of-war reports, faxed in from the Saudis, referenced capturing several journalists. I said, "I need to talk to this guy."

Did you take notes during interpreting?

Yes, I wrote down almost every word I could. I tried to take down the entire conversation in my notes. Although we had the tapes, I still wanted my own notes because we had to type them up.

During your lecture at the Institute, you mentioned that sometimes you wrote down additional information for the U.S. generals.

Occasionally I did for General Johnson. At Safwan, for example, the Iraqis were responding to our expressed concern about the use of helicopter gunships against the Shi'a rebels in the south, whom we could see from our positions, yet we had given the permission to fly helicopters at the earlier meeting. In response to General Johnson's concern, the Iraqi military intelligence general went on this long diatribe of how the Iranians have infiltrated the Revolutionary Guard Corps into southern Iraq and were fomenting revolution. I would just lean over and write,

1494

party line. After I gave General Johnson the gist of it, he just cut him off by saying, "It doesn't matter."

Did you finish unfinished sentences for the speaker?

No, never. I don't recall anyone ever finishing a sentence for an Iraqi or a Saudi officer. Occasionally they would say, "You know what I mean." and I would ask him to say it again. I didn't want to be in the position of trying to put someone else's thoughts, especially of an Iraqi or of a Saudi, into words unless I had their words.

Would they sometimes repeat themselves?

Yes, they would repeat utterances frequently. If they were giving an answer to a complex question, they would start, stop, and start again. During normal conversation, I would paraphrase it, or gist it, to get the meaning across.

Did they also ask you for clarifications if they could not understand what you were saying?

Yes. Once in a while, they would come to us and say "We don't understand this phrase. What did you mean?" They would have several translations of an English-language document into Arabic. They would say, "We think it means *this*" or "it could mean *this*" or "it could mean *that*" and we had to pick the one that best conveyed the meaning. That was rare, because when we wrote in English to the Saudis, and they did the same when they wrote in Arabic, we tried to be clear. I always recommended to the CENTCOM writers handling the correspondence to be to the point and to make sure it could be easily translated into Arabic. I also asked them not to use esoteric terms or abstract references. For the most part, the Saudis expressed themselves clearly to us.

Occasionally, I would take a phrase to another one of the American interpreters and say, "How would you translate this phrase? What does it mean to you" Most of the time he would respond, "It means *this* to me." I said, "I know, but in the context of the letter, this rendition does not make sense. Why would they say that?" Many times we would go to the officer that wrote it and say we were translating this text for the general. We would ask him to paraphrase it or to clear it up. That was the beauty of working with the allies. It was a little harder to clarify communications with the Iraqis.

1495

If a speaker, an Iraqi, for example, made an error during interpretation would you let them know?

I was concerned about what their interpreter was telling the Iraqi general. I would listen to his interpretation. If it was wrong, and I thought he was getting the wrong message, I would ask to talk to the interpreter. I would try to do it in a break because I did not want to disrupt the talks. I would say "Excuse me. I think we have a problem with the translation." or, politely, "I'm not sure I understood what you said," and ask him to explain it.

Occasionally they would come to us and say, "We don't think you understood the message." The Iraqi interpreters that I dealt with were excellent speakers of English.

What would happen if you noticed an obvious mistake?

This came up during the initial talks at Safwan. The Iraqis would speak to us in Arabic and we would interpret for General Schwarzkopf into English. He would speak English and their interpreter would translate it into Arabic: always into your native language. We always checked each other's interpretation, that's one of our jobs. We had three interpreters on the U.S. side. We all noticed that the Iraqi interpreter was translating the word for *prisoners of war* and *detainees*, primarily the Kuwaiti detainees, as *guests*. We thought this was an improper interpretation, because in Arabic the word for *prisoner of war*, a military person, is very specific. They were using the right equivalent when we were talking about the U.S. and coalition prisoners, but when we were talking about the Kuwaiti civilians that had been arrested and taken to Iraq, they were using the equivalent of *guests*. We didn't disrupt the flow of the conversation, but at one of the breaks, we went over to talk to the Iraqi interpreter and we explained to him that we felt that they were using the wrong term, and we would like to correct the record. We gave him the word we preferred and they came back with a different one that we both agreed was okay. That was one of those situations where we just interjected ourselves, but we decided to intervene during a break, not the actual back-and-forth exchange between the two principals.

I noticed that Saddam Hussein also calls hostages guests *in his speeches.*

He always has. They did not use the equivalent of *hostage*. We found a word that meant *detainees* and they agreed to use it. We didn't like the word *hostage* either.

1496

This exchange illustrates how sensitive the use of words can be.

Since General Schwarzkopf did not understand Arabic, this improper translation of the term did not disrupt what he was doing. The senior Iraqi was not concerned about it, either. But for the historical record, we wanted it to be correct. General Khalid on the Saudi side was a little upset with that translation as well, because he understands both English and Arabic perfectly. We also told his staff that we were correcting the term and he said "Absolutely. I agree with it 100%."

This exchange was conducted in Modern Standard Arabic, not Iraqi?

Iraqi interpreter was translating from English to Arabic in a mix of "Iraqi-accented" Standard. I think too much is made of the dialects. I have found that with a solid proficiency in Modern Standard Arabic, you can go into any country and pick up enough of the local slang, vocabulary, and pronunciation to make yourself understood.

It is always great to be trained in the dialect, but sometimes we do not have that opportunity. Although I was trained at the Institute in Modern Standard with a little bit of Egyptian dialect, I never served in Egypt. We don't always have the luxury of specializing down to the dialect level in the U.S. forces. Our missions take us to Jordan, Egypt, North Africa, Saudi Arabia, and the Gulf states. From a management perspective, to train people in the Iraqi dialect only is not a good idea. Thus, Modern Standard is probably the most important tool. Modern Standard allowed me to go to any other country and pick up the dialect.

Have you experienced any lapses of attention?

Sometimes I experienced lapses of attention due to the content of a particular passage or a particular meaning. While I was figuring out meaning, he was still talking. By the time I realized the meaning, I missed part of his turn. I hoped that I could pick it up from context later.

I'm sure, occasionally, I missed things: It is just the nature of the game.

What signs of overload did you experience?

When I could not immediately pick up on what was going on. Every time that has happened to me, I have always had somebody else there. We would work out little signals to each other, usually with Colonel

1497

Dunn or a Kuwaiti student that worked with us. I would just nod to them and they would pick it up.

Many of the students at the Institute would like to become interpreters. What does it take in terms of education, training, and personality to become an interpreter?

Assuming that you have mastered the required level of language, you have to understand the culture. I'm not talking about the ancient culture, arts and sciences, but about the background of these people and their recent history. If you are aware of the major issues of the Middle East, particularly following the breakup of the Ottoman Empire, you have a solid background in what has shaped the thinking of the people you are going to be dealing with. You have to like people. You cannot function as an interpreter if you are not comfortable talking to strangers at length. Also, when you are acting as an interpreter, you have to realize that you are a tool of your principal. You are not the negotiator, you are the interpreter. A lot of times there is an instinct to just cut through a problem, but you have to realize you are there as a facilitator, and not the actual conductor.

You are not supposed to step into the action, right?

It is very difficult not to do that; it is very tempting. At times, of course, everybody does it. If it is something small, it is not a problem. When General Schwarzkopf wanted to make sure there was no question about his words, he would remind me to translate exactly what he was saying.

Let us focus on training. Although job experience forges the best schools, the "tuition" can be priceless and the final outcomes irrevocable. What lessons can you transfer into schools such as DLIFLC where curriculum can be planned, time made available, and the cost controlled?

How do you train someone to the level of ancillary skills required of an interpreter? You can provide the foundation at the school. The language, of course, is the key. You have to have solid language skills to be an interpreter. There's no way around that. But that's not enough. You have to understand the history of the people you are dealing with. You have to understand also the military because you have to deal with the military situation. If you are working in a political environment, such as an embassy, like we did in Damascus, you have to understand the political situation. You have to know your country's interests and, more im-

1498

portantly, your host country's interests—where they meet and where they diverge. I don't know how you replicate it in a training environment.

My experience and education prepared me for interpreting assignments. I was selected to be an interpreter in 1990, and that was the first time I did any serious interpreting. Earlier I was teaching. I had learned Arabic and worked in Arabic since graduating in 1974. I had served in Arab countries and traveled extensively in the Arab world. I had worked military and political issues in the Arab countries, and served as an adviser in a Jordanian unit. This experience gave me a good understanding of the Arab military culture.

Regional studies are very important. When I was at the Institute in 1974, we had interesting history books of each country in the language. The talk I'm giving brings that sort of thing to the students. Understanding Iraq is not reading the history of Iraq. You have to look at what happened to them. The defining characteristics of Iraq took place from the end of World War I until now. The problem with that approach is that to do it with every country that speaks Arabic would be a daunting task.

How did your training at DLIFLC prepare you for your work? What would you add or change in it to prepare today's students for their future duties?

Most of the DLIFLC students go into the Cryptologic Training System. While training and supervising graduates of the Institute, I noticed a decline of language capability in the 1980s and then an increase from the late 1980s through now. The lengthening of the program (from 47 to 63 weeks starting in 1992) had a lot to do with the improvement. I would have loved to have had 16 months of training. I say that now: I did not have to sit here for 16 months.

The schools do a good job. Even if you do not graduate speaking fluently, or understanding everything, the grammatical foundation you get in Modern Standard is good. I've served with people who have graduated from both civilian universities that have taught them Arabic and I have served with people that have gone to the State Department's Foreign Service Institute and also to contract schools. Invariably, I find that the best linguists are those that are trained at DLIFLC that have the opportunity to live in an Arabic-speaking country.

Major Emily Francona comments: Having the basics honed in country is probably the best combination you could have. I was not fortunate enough to learn any of my languages at DLIFLC. Comparing

the curriculum and the results, the Institute is definitely far superior. I would have preferred to have learned my languages at the Institute.

Any particular activities would you recommend that would help in performance of a military interpreter? You mentioned frequently you were handling routine tasks.

Translation. The interpreting was the most important task we did, but translation was the most frequent. Written documents would come to us, they were so diverse that we never knew what the subject would be. Even though they came from one military officer to another, the subject could be the *no smoking policy* in the headquarters, office assignments, communication schedules, access to prisoners, intelligence updates, or situation reports.

Could you tell us how good interpreters can prepare themselves for the first assignment?

Learn as much of the history as possible. If you are going to be called on to interpret, learn as much of the subject matter as you can, so that when you hear it, it is not a surprise to you. First, if you know the subjects that are going to come up, make sure you know the vocabulary. In one of the situations, we were going to talk about repatriation of prisoners of war. So we anticipated terms and phrases pertaining to logistics and timetables. Refresh your memory to make sure you are aware of the words the Iraqis or the Saudis may use. Second, look at a map. Familiarize yourself with the geography. Then when they mention the name of a town, you are not searching around for it on a map: You know where it is located and how it is spelled. Frequently, the principal will not only ask you for the geographical name, but also for its location. You are expected to know these basics. In summary, the more ancillary knowledge you have about the subject, the less stressful the interpreting is going to be.

In a way, you have to project into the future what will happen and what you will need when it happens. From the hind sight of your experience, what else would you do in preparation for interpreting?

I would have, probably, learned more of the Saudi dialect. When I went to Saudi Arabia, I knew I was going to be General Schwarzkopf's interpreter. I assumed correctly that I would be talking to Iraqis or debriefing Iraqi prisoners of war. The majority of the time, however, I talked to the

150

Saudis in Modern Standard. Although it was adequate, I wish I could have communicated with them in their dialect.

There are numerous training aids such as dialect textbooks. You say "I'm going to Saudi Arabia." Do I need Saudi dialect? Do I need Iraqi dialect? Who am I going to be talking to most? I felt comfortable with the Iraqis and the Saudis, but I could have been a little more effective, on an interpersonal basis, had I spoken a little more Saudi dialect. We had one interpreter assigned to us, an army lieutenant colonel reservist, who had gone to the Saudi Command College and was just wonderful in Saudi dialect. A little slow in the Iraqi dialect, but great in Saudi. Because of his proficiency in this dialect, his rapport with the Saudis was great. I found that rapport is important if you are going to be interpreting for the same people all the time, or briefing the same people. Good rapport makes things easier.

In other words, you have to, partially, fit a person to the task.

Absolutely. We rarely talked to the same Iraqi twice. We would do our job, and we probably would never see them again. In the headquarters, I saw the same Saudi officers everyday; twice a day sometimes. For example, I got to know General Madani on a social and professional level because I briefed him in Arabic twice a day.

What enhancement routines, that you have used, would you recommend for our graduates?

I kept current with radio and print media. Today, Arabic satellite, cable programming, an expanded VTT (Video Tele Training) system and internet content, provide excellent sources for language enhancement. All these tools have great potential for one-on-one tailored training, as well as language maintenance and enhancement.

Thank you very much for your advice to our students. I wish you and your wife success in your future endeavors.

Notes

[1] The Conference on *Tradition and Innovation in Translation and Interpretation* took place in February 1999 at the Monterey Institute of International Studies.

[2] N. Goroshko & L. Slutsky. (1993). Four-Handed Teaching. *Dialog on Language Instruction, 9*/1, 49-53.

[3] H. Mikkelson. (1998). Towards a Redefinition of the Role of the Court Interpreter. *Interpreting: International Journal of Research and Practice in Interpreting, 3/2 42.*

1501

Reviews

Teaching by Chatting. (1998). By J. M. Cots. Lleida: Universitat de Lleida. Pp. 205, paper, ISBN 84-89727-66-X.

Reviewed by LEO VAN LIER
Monterey Institute of International Studies

This book, the first monograph in a series called *Quadernas de Sintagma* (*Sintagma* is a journal published by the University of Lleida), is subtitled "A Pragmatic Analysis of Instructor-Student Conversations at an American University." It provides a detailed discussion of what is called "office hours" at universities, a topic that has not been extensively investigated, even though it provides an important academic context for university students. The corpus of data consists of 20 office-hour conversations, ranging from less than two minutes to over 40 minutes.

The study is an interesting exercise in the analysis of spoken interaction from several different perspectives. But those who are looking for explicit advice on the office hour, or for a critical perspective on the professor–student relationship as developing outside of lectures and classrooms, will be disappointed. The study is an explicitly neutral discussion of office-hour discourse, and the reader will have to decide what to think of it in academic and pedagogical terms. The analysis is like a mirror: Ultimately readers have to decide what it is that they see, and how they like it.

The first part of the book is a general introduction to communicative and pragmatic competence. Apart from a brief introductory "walking tour" in which role plays are used to show how non-native and native speakers participate in academic communicative encounters, this section is not explicitly related to the topic of office hours. Rather, it is an overview of various theories and models of spoken interaction. This section, consisting of about fifty pages, is a very lucid and well-organized overview of various models of communicative competence and of spoken interaction. As an introduction to these topics, it is competently and efficiently done, and worth reading for its own sake, quite apart from any interest the reader may have in the office-hour encounter itself.

The second and main part of the book consists of an

1502

ethnographic description of instructor-student interaction during office hours. In this description the various models and theories from the first part of the book are used to elucidate samples of instructor–student interaction. As Cots points out, the office hour encounter falls between the institutional and the personal areas of social action, more so than the classroom or lecture hall.

Cots uses the notion of self-presentation to illustrate strategies such as "avoiding assertiveness," "explaining and justifying," "displaying a positive self," "emphasizing modesty," and "showing attitudes and feelings." He also discusses the negotiation of formality, power and distance, and the minimization of imposition.

A separate chapter deals with discourse competence by analyzing topic management, turn taking, and information structure. Repairing is discussed under the heading of strategic competence, although the boundary between discourse competence and strategic competence seems rather blurred. The final chapter shows the various aspects of analysis in one coherent whole by analyzing one complete encounter (albeit a short one) using all the categories and strategies explained in the book. The appendix contains transcripts of three further encounters that could be used by readers or students of conversation analysis to conduct similar analyses.

The book contains a thorough overview of the analysis of spoken discourse in the context on one type of speech event, the office hour encounter at an American university. It integrates four approaches to descriptive pragmatics: sociolinguistics, conversation analysis, discourse analysis, and the ethnography of speaking. It serves as a very useful text on descriptive pragmatics and sheds interesting light on an important but not much studied speech event. It deserves to be widely read by students of sociolinguistics and pragmatics, as well as by those who hold office hours.

1503

Language and Development: Teachers in a Changing World. (1997). Editors: B. Kenny and W. Savage. (*Applied Linguistics and Language Study Series*. General Editor: *Christopher N. Candlin*.) New York: Addison Wesley Longman.*

Reviewed By KEVIN W. K. CHU
City University of Hong Kong

This collection of selected papers was originally presented at the Conference of the Regional Education Language Center (RELC) on Language Programs in Development Projects in April 1993 at the Asian Institute of Technology in Bangkok. The papers document English Language Teaching in Cambodia, China, India, Indonesia, Laos, Malaysia, Papua, New Guinea, the Philippines, Singapore, and Vietnam. The book illuminates "short stories of teachers' professional lives in a changing world" rather than scholarly written papers (p. 4). The authors of **Language and Development: Teachers in a Changing World** (hereafter **L and D**) discuss recent education policy, English for Special Purposes (ESP) issues, and foreign aid for foreign language education in these countries. The book is organized into three parts titled *Coping with Change*, *Teaching and Learning in Different Worlds*, and *Responding to the Players*; and consists of 21 chapters plus introductory and concluding chapters.

The first seven chapters in **L and D** demonstrate how practitioners "cope with change" by reconstructing education programs, introducing staff development programs, by developing teacher support teams, identifying students' purposes of learning English, and by using task-based activities such as modified role play. The papers focus on practical concerns such as large-class size, low motivation, lack of budget and staff, and time constraints.

Teaching and Learning in Different Worlds is also composed of seven chapters which document cases of adapting culturally incompatible teaching materials and also cases of integrating language teaching with mainstream subjects in the forms of simulation exercises and business visit projects. In order to implement teacher education reforms, either in classes of trainee pilots, engineering students or electrical technicians, language teachers, educators, and development professionals struggle with traditional beliefs, bureaucracy, and fear of change. The chapters exemplify well how reflective the practitioners are in making these adaptations to "different worlds."

The area of the conflicting objectives and needs of the foreign language education programs is discussed in *Responding to the Play-*

1504

ers. By the *Players*, Kenny and Savage refer to those who deliver and receive foreign aid for foreign language education. Some telling stories in this section will initiate discussion of the roles and expectations of funding agencies, managing agencies, language teaching consultants, politicians, academics, and teachers.

I appreciated the clear account of characteristics in the concluding chapter of **L and D**. Here, Savage admirably justifies the new emerging field of language and development by arriving at its five notable characteristics. Namely, they are: (1) change-oriented, (2) experiential, (3) pro-autonomous, (4) collaborative, and (5) communicative. This interpretative commentary is a good beginning of a discussion forum for language planning issues, linguistic imperialism, and language rights that have not been made explicit in these "developing" societies. I hope that future language-and-development experts will expand on these issues at length. Nonetheless, **L and D** is extremely valuable in presenting the development of language training; in establishing the forum for the voices of the local teachers and expatriate language educators; in raising awareness of the effect of the cultural, economic and sociopolitical factors on language policies; and, finally, in providing practical suggestions for teachers. In this regard, the stories present the readers with concrete images of how to implement a critical pedagogy—how efficient the players might be in applying professional knowledge; how sensitive to the institutional and social contexts they ought to be; and how, finally, they could not simply accept things as they were.

We can see the enthusiasm of **L and D** to show us around "the changing world." The chapters can be read properly only in the light of full appreciation of the difficulties confronted by the players, the frontline teachers and, most importantly, the course participants. I have no doubts that readers could arrive at their own meanings, as invited by Kenny and Savage in the introductory chapter titled "Setting the Scene," and could reflect upon the experience shared by the contributors. How are their practices constrained and influenced by institutional, social, historical, ethical, and political factors? By unveiling a wider scope of teaching contexts, this book helps readers question taken-for-granted and unarticulated assumptions and consequently alter their perspectives on language study and language education.

*This is a substantially expanded version of the book review published in TESL-HK, a newsletter for ELT professionals in Hong Kong, June 1998, 2.

1505

Making Communicative Language Teaching Happen: Directions for Language Learning and Teaching. (1995). By James F. Lee and Bill VanPatten. New York: McGraw-Hill.

Reviewed by PHILIP A. WHITE
Defense Language Institute Foreign Language Center

Many teachers of foreign languages may have come to the field from different ethnic and professional groups. Some learned the language they teach at an institution, while others grew up with it at home. Especially the latter may want to gain a better understanding of the field of foreign language education and contemporary methods. If so, then this book, *Making Communicative Language Teaching Happen*, could served as a good start.

As the title indicates, this text provides a basic introduction to the approach of communicative language teaching. The two authors are well prepared to do so as both are widely known within the field of foreign language education. Thus, the reader should have some assurance of the usefulness and "authenticity" of the material.

The authors have written this text as a "guide to helping instructors develop a communicative classroom environment that blends listening, speaking, and writing (p. ix)." Their audience is those who are graduate students (the book is planned as a text for a course having been "field tested" with the authors' students) and practicing teachers who need a resource manual for developing tasks and material for the classroom. They note that there over 200 activities and test sections included (p. x).

Topics include new roles for teachers and students, the importance of "comprehensible mean-bearing input in second language acquisition," an approach to grammar instruction, classroom oral communication with suggestions, contemporary approaches to teaching reading and writing, suggestions on building towards proficiency, and testing principles. The book contains suggestions and questions for reflection throughout the book, suggestions for further reading, and end-of-chapter activities (presumably for classroom and students using the text in a course), and an associated workbook for the text (not reviewed here).

The book contains 13 chapters and an epilogue, organized into five sections of varying size, a bibliography, and an index. The five major sections of the text are: (1) Preliminary considerations, (2) Grammar instruction, (3) Spoken language, (4) Reading and writing and (5) A look forward.

The material the authors provide is contemporary and reflective of the field as well as indicative of the approach they use. In some areas perhaps the reader is offered too much material that might be of minimal interest to the teacher or persons now in preparation for teaching. Particularly their discussion of the older, and now generally unused audiolingual method, is more defensive in nature than informative and useful for the classroom (p. 7).

After establishing a view of a classroom teacher that the two authors consider common in teaching, that of the instructor as an "Atlas figure" holding the entire weight of the classroom and learners on his shoulders, they offer a countervailing perception of the teacher as an architect and resource person within the class (chapters 1 and 2). Unfortunately, they then proceed to equate the instructor-as-Atlas view with audiolingualism. While the view of the teacher as a facilitor and as a resource has much to recommend it, the defensive attacks on audiolingualism neither offer much to the reader nor enhance the text.

Lee and VanPatten provide a detailed and useful view of teaching grammar within the communicative language teaching approach. The debate over the place of grammar in the foreign language classroom has been a historic one within the field and is not limited to any one approach. The authors come down on the side of teaching grammar and doing so through "structured" input and output in the classroom. Rather than have learners perform activities that force manipulation of their output, the view offered here is that of structured input to offer the learner an opportunity for meaning-bearing activities. In support of this approach to grammar, they offer a wealth of activities for input and output (chapters 5 and 6).

Readers who find the presentation of grammar at least somewhat alien to a traditional view should be aware that the two authors are admittedly influenced by the work of S. Pit Corder, a British educator who is closely associated with the concept of "code switching" within the field of sociolinguistics as well as error analysis. At the beginning, they cite Corder's view that language might be an activity that cannot be taught, but can only come about within conditions conducive to acquisition by the learner (p. 35).

Readers who are seeking a contemporary means of handling pronunciation in the classroom will be disappointed by the writings. While the text contains a discussion of listening comprehension with appropriate exercises, pronunciation does not even appear in the index. As most people, whether teachers or students, are aware that normally human language involves the production of sound and that

1507

the sounds of another language are rarely identical to their native speech, the absence of any discussion of pronunciation is puzzling. In a text claimed as being developed for practicing teachers, some sort of statement about pronunciation seems obviously necessary—why are or aren't activities specifically for pronunciation necessary? How is the learner supposed to make him or herself comprehensible to the native speaker? On these questions as well, the authors are silent.

As noted, the text contains many activities for the classroom. From the perspective of those who teach languages located outside of western Europe, the restriction of examples to primarily Spanish could be viewed as a limiting factor. Of course, as the authors hold positions in university Spanish departments this limitation perhaps is understandable.

All in all, in spite of its faults—and what book is flawless?—**Making Communicative Language Teaching Happen** is a text that can be read with much profit by most classroom teachers. Most readers will benefit not only from the example activities, but also from the extensive suggestions for reading given at the end of each chapter.

1508

The Road Ahead. (1995). By William H. "Bill" Gates, III. New York: Viking. Pp. 286, CD version included, ISBN 0-670-77289-5.

Reviewed By RODERIC A. GALE
Defense Language Institute Foreign Language Center

"First and foremost, Bill Gates is an idea man," said Barbara Walters on December 22, 1998 during the ABC News Special *The 10 Most Fascinating People of 1998*. This characterization clearly reflects the message coming from *The Road Ahead*.

In keeping with everything one expects of the Microsoft Mastermind, this book is always looking forward, with just enough history to set the context for that look and to prove the idea that looking forward is the only way to go, whether in dealing with information systems or anything else. Selections from the Contents page give insight to that direction: A Revolution Begins, Lessons from the Computer Industry, Applications and Appliances, Implications for Business, Education: The Best Investment, and Critical Issues.

In reviewing this book for the *Applied Language Learning*, I have looked at it from the perspective of what it has to offer for those involved in language teaching and evaluation. However, before addressing that assessment, an overview of the book is warranted. Bill Gates did not write this, his first publication, as a historical review of his life, of the software business, or his contributions to that business. Those factors do appear throughout the book, but only to provide the setting to look forward and the means by which events have moved forward. It is interesting that the perhaps richest person in the world has not focused on what he has done, but rather on what has been done and the lessons to be drawn from his work. His thoughts focus on what the future holds in store and how we can either seize the opportunities ahead or be held captive to them.

Just as he speaks in public or television appearances, Gates writes with a casual approach that makes for a fast read. It is not at all what one may expect from someone who virtually created the high tech software world which impacts everyone around the world in one way or another. There is no technical information to be gleaned and very few notes to be taken on the ideas he presents. As a result, it can be read in a weekend or over a week of evenings. The compartmentalization of the chapters makes it easy to move through the book at will.

What of the chapter that most applies to language education though? "Some fear that technology will dehumanize the formal

1509

educator.... But. . . technology can humanize the education environment. Corporations (read, teaching organizations) are reinventing themselves around the flexible opportunities afforded by information technology, classrooms will have to change as well. (p. 184) There is an often-expressed fear that technology will replace teachers. I can say emphatically and unequivocally, IT WON'T. However, technology will be pivotal in the future role of teachers" (p. 185). I believe one can say the same for testers and evaluators of students. As a result, the full spectrum of people in the foreign language process—students, teachers, and testers—should be challenged and assured by his comments.

Clearly, this man Bill Gates, who may arguably be the person most influential in putting computer software in the home and classroom, is calling for those in education to leverage the use of computers in their teaching. He adamantly rejects the thought that machines and programs will replace these educators. "Educators. . . are, among other things, facilitators. . . they will have to adapt and readapt to changing conditions. Unlike some professions, however, the future of teaching looks extremely bright" (p. 187).

The Road Ahead chapters on Education: The Best Investment and Critical Issues are part of the spectrum of light along the tunnel of education. We in that profession may do well to rally behind this White Knight.

1510

General Information

Calendar of Events*

1999

6-9 March, *American Association of Applied Linguistics*, Stamford. Information AAAL, (612) 953-0805, Fax (612) 431-8404, PO Box 21686, Eagan, MN 55121-0686; Email [aaaloffice@aaal.org].

8-14 March, *Teachers of English to Speakers of Other Languages*, New York. Information TESOL, (703) 836-0774, Fax (703) 836-7864, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; Email [conv@tesol.edu], URL [www.tesol.edu].

11-13 March, *Southern Conference on Language Teaching with Foreign Language Association of Virginia*, Virginia Beach. Information Lynne McClendon, SCOLT Executive Director, (770) 992-1256, 165 Lazy Laurel Chase, Roswell, GA 30076; Email [lynnemcc@mindspring.com].

7-10 April, *Pacific Northwest Council for Languages*, Tacoma. Information PNCFL, PO Box 4649, Portland, OR 97208-4649; Email [112063.622@compuserve.com].

8-11 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, Dickinson College, (717) 245-1977, Fax (717) 245-1976, PO Box 1773, Carlisle, PA 17013-2896; Email [nectfl@dickinson.edu], URL [www.dickinson.edu/nectfl].

8-11 April, *American Hungarian Educators Association*, Cleveland. Information Martha Pereszlenyi-Pinter, Classical & Modern Languages & Cultures, (216) 397-4723, FAX (216) 397-4256, John Carroll University, Cleveland, Ohio 44118; Email [mpereszlenyi@jcvaxa.jcu.edu], URL [http://www.magyar.org/home.html].

15-18 April, *Central States Conference on the Teaching of Foreign Languages*, Little Rock. Information CSCTFL, Rosalie Cheatham, (501) 569-8159, Fax (501) 569-8157, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; Email [rmcheatham@ualr.edu].

*Courtesy of *The Modern Language Journal* (University of Wisconsin)

1511

13-15 May, *JNCL-NCLIS Delegate Assembly*, Washington. Information JNCL-NCLIS, (202) 966-8477, 4646 40th St. NW, Third Floor, Washington DC 20016; Email [info@languagepolicy.org], URL [http://www.languagepolicy.org].

20-23 May, *Language Teacher Education*, Minneapolis. Informa tion International Conference on Language Teacher Educa tion, CARLA, (612) 627-1870, Fax (612) 624-1875, UTEC, Suite 111, 1313 5th St SE, Minneapolis, MN 55414; E-mail [carla@tc.umn.edu], URL [http://carla.acad.umn.edu].

22-30 May, *Conseil International d'Etudes Francophones*, Lafayette. Information Ginette Adamson, Fax (316) 978-3319, Modern Languages, Wichita State University, Wichita, KS 67260-0011; Email [adamson@twsuvm.uc.twsu.edu].

30 May-4 June, *Computer Assisted Language Instruction Consortium*, Oxford. Information Esther Callais, (512) 245-1417, Department of Modern Languages, Southwest Texas State University, San Marcos, TX 78666; Email [info@calico.org].

3-6 June, *ADFL Seminar West*, Palo Alto. Information Association of Departments of Foreign Languages, Attn: Elizabeth Welles, 10 Astor Place, New York, NY 10003-6981; Email [elizabeth.welles@mla.org].

11-14 July, *American Association of Teachers of French*, St. Louis. Information AATF, (618) 453-5731, Fax (618) 453-5733, Mailcode 4510, Department of Foreign Languages, Southern Illinois University, Carbondale, IL 62901-4510; Email [abrate@siu.edu], URL [aatf.utsa.edu].

30 July-3 August, *American Association of Teachers of Spanish & Portuguese*, Denver. Information AATSP, (970) 351-1090, Fax (970) 351-1095, Butler-Hancock Hall #210, University of Northern Colorado, Greeley, CO 80639; Email [lsandste@bentley.unco.edu].

16-21 August, *International Association of Teachers of Russian Language and Literature*, Bratislava (Slovakia). Information American Council of Teachers of Russian (ACTR), (202) 833-7522, Fax (202) 833-7523, 1776 Massachusetts Ave. NW, Suite 700, Washington, DC 20036; Email [ddavidson@actr.org].

28-30 October, *Foreign Language Association of North Carolina*, High Point. Information Debra S. Martin, FLANC Executive

1512

Director, (828) 686-4985, Fax (828) 686-3600, PO Box 19153, Asheville, NC 28815; Email [martintl@interpath.com].

4-6 November, *Wisconsin Association of Foreign Language Teachers*, Appleton. Information Kyle Gorden, (414) 723-6316, 4969 Hickory Court, Elkhorn, WI 53121; Email [kylegorden@elknet.net].

17-18 November, *National Association of District Supervisors of Foreign Languages*, Dallas. Information Sharon Watts, (402) 557-2440, Omaha Public Schools, 3215 Cuming, Omaha, NE 63131; Email [swatts@ops.org].

18-21 November, *American Association for the Advancement of Slavic Studies*, St. Louis. Information AAASS; Email [walker@core-mail.fas.harvard.edu].

19-21 November, *American Council on the Teaching of Foreign Languages*, Dallas. Information ACTFL, (914) 963-8830, Fax (914) 963-1275, 6 Executive Plaza, Yonkers, NY 10701-6801; Email [actflhq@aol.com], URL [http://www.actfl.org].

19-21 November, *American Association of Teachers of German*, Dallas. Information AATG, (609) 795-5553, Fax (609) 795-9398, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America*, Chicago. Information MLA, Fax (212) 477-9863, 10 Astor Place, New York, NY 10003-6981; Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, Chicago. Information Masako Ueda, (401) 863-3933, Fax (401) 863-7330, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; Email [masako_ueda@brown.edu].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, Chicago. Information AATSEEL, Fax (520) 885-2663, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

2000

24-26 February, *Southern Conference on Language Teaching with Alabama Association of Foreign Language Teachers*,

Birmingham. Information Lynne McClendon, SCOLT Executive Director, (770) 992-1256, 165 Lazy Laurel Chase, Roswell GA 30076; Email [lynnemcc@mindspring.com].

10-13 March, *Central States Conference on the Teaching of Foreign Languages*, TBA. Information CSCTFL, Rosalie Cheatham, (501) 569-8159, Fax (501) 569-8157, University of Arkansas - Little Rock, 2801 S. University Avenue, Little Rock, AR 72204; Email [rmcheatham@ualr.edu].

11-14 March, *American Association of Applied Linguistics*, Vancouver. Information AAAL, (612) 953-0805, Fax (612) 431-8404, PO Box 21686, Eagan, MN 55121- 0686; Email [aaaloffice@aaal.org].

14-18 March, *Teachers of English to Speakers of Other Languages*, Vancouver. Information TESOL, (703) 836-0774, Fax (703) 836-7864, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; Email [conv@tesol.edu], URL [www.tesol.edu].

13-15 April, *Pacific Northwest Council for Languages*, Missoula. Information PNCFL, PO Box 4649, Portland, OR 97208-4649; Email [112063.622@compuserve.com].

13-16 April, *Northeast Conference on the Teaching of Foreign Languages*, Washington DC. Information Northeast Conference, (717) 245-1977, Fax (717) 245-1976, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; Email [nectfl@dickinson.edu], URL [www.dickinson.edu/nectfl].

4-6 May, *Balkan and South Slavic Linguistics, Literature and Folkore*, Lawrence. Information Marc L. Greenberg, Dept. of Slavic Languages and Literatures, Fax (785) 864-4298, 2134 Wescoe Hall, Lawrence, KS 66045-2174; Email [m-greenberg@ukans.edu].

TBA July, *American Association of Teachers of French*, Paris. Information AATF, (618) 453-5731, Fax (618) 453-5733, Mailcode 4510, Department of Foreign Languages, Southern Illinois University, Carbondale, IL 62901-4510; Email [abrate@siu.edu].

1-5 August, *American Association of Teachers of Spanish & Portuguese*, San Juan. Information AATSP, (970) 351-1090, Fax (970) 351-1095, Butler-Hancock Hall #210, University of Northern Colorado, Greeley, CO 80639; Email [lsandste@bentley.unco.edu].

2-4 November, *Foreign Language Association of North Carolina*, High Point. Information Debra S. Martin, FLANC Executive

Director, (828) 686-4985, Fax (828) 686-3600, PO Box 19153, Asheville, NC 28815; Email [martintl@interpath.com].

9-12 November, *American Association for the Advancement of Slavic Studies*, Denver. Information AAASS; Email [walker@core-mail.fas.harvard.edu].

15-16 November, *National Association of District Supervisors of Foreign Languages*, Boston. Information Sharon Watts, (402) 557-2440, Omaha Public Schools, 3215 Cuming, Omaha, NE 63131; Email [swatts@ops.org].

17-19 November, *American Council on the Teaching of Foreign Languages*, Boston. Information ACTFL, (914) 963-8830, Fax (914) 963-1275, 6 Executive Plaza, Yonkers, NY 10701-6801; Email [actflhq@aol.com], URL [http://www.actfl.org].

17-19 November, *American Association of Teachers of German*, Boston. Information AATG, (609) 795-5553, Fax (609) 795-9398, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America*, Washington, D.C. Information MLA, Fax (212) 477-9863, 10 Astor Place, New York, NY 10003-6981; Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, Washington, D.C. Information Masako Ueda, (401) 863-3933, Fax (401) 863-7330, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; Email [masako_ueda@brown.edu].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, Washington, D.C. Information AATSEEL, Fax (520)885-2663, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

2001

24-27 February, *American Association of Applied Linguistics*, St. Louis. Information AAAL, (612) 953-0805, Fax (612) 431-8404, PO Box 21686, Eagan, MN 55121-0686; Email [aaaloffice@aaal.org].

27 February-3 March, *Teachers of English to Speakers of Other Languages*, St. Louis. Information TESOL, (703) 836-0774,

Fax (703) 836-7864, 1600 Cameron St., Suite 300, Alexandria, VA 22314-2751; Email [conv@tesol.edu], URL [www.tesol.edu].

19-22 April, *Northeast Conference on the Teaching of Foreign Languages*, New York. Information Northeast Conference, (717) 245-1977, Fax (717) 245-1976, Dickinson College, PO Box 1773, Carlisle, PA 17013-2896; Email [nectfl@dickinson.edu], URL [www.dickinson.edu/nectfl].

14-15 November, *National Association of District Supervisors of Foreign Languages*, Washington, D.C. Information Sharon Watts, (402) 557-2440, Omaha Public Schools, 3215 Cuming, Omaha, NE 63131; Email [swatts@ops.org].

16-18 November, *American Council on the Teaching of Foreign Languages*, Washington, D.C. Information ACTFL, (914) 963-8830, Fax (914) 963-1275, 6 Executive Plaza, Yonkers, NY 10701-6801 Email [actflhq@aol.com], URL [http://www.actfl.org].

16-18 November, *American Association of Teachers of German*, Washington, D.C. Information AATG, (609) 795-5553, Fax (609) 795-9398, 112 Haddontowne Court #104, Cherry Hill, NJ 08034; Email [73740.3231@compuserve.com].

27-30 December, *Modern Language Association of America*, TBA. Information MLA, Fax (212) 477-9863, 10 Astor Place, New York, NY 10003-6981; Email [convention@mla.org].

27-30 December, *North American Association of Teachers of Czech*, TBA. Information Masako Ueda, (401) 863-3933, Fax (401) 863-7330, Box E, Department of Slavic Languages, Brown University, Providence, RI 02912; Email [masako_ueda@brown.edu].

27-30 December, *American Association of Teachers of Slavic & E. European Languages*, TBA. Information AATSEEL, Fax (520) 885-2663, 1933 N. Fountain Park Dr., Tucson, AZ 85715; Email [76703.2063@compuserve.com], URL [http://clover.slavic.pitt.edu/~aatseel/].

1516

Information for Contributors

Statement of Purpose

The purpose of *Applied Language Learning* (*ALL*) is to increase and promote professional communication within the Defense Language Program and academic communities on adult language learning for functional purposes.

Submission of Manuscripts

The Editor encourages the submission of research and review manuscripts from such disciplines as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields such as linguistics, education, communication, psychology, and social sciences; (5) assessment of needs within the profession.

Research Article

Divide your manuscript into the following sections:

- Abstract
- Introduction
- Method
- Results
- Discussion
- Conclusion
- Appendices
- Notes
- References
- Acknowledgements
- Author

Abstract
Identify the purpose of the article, provide an overview of the content, and suggest findings in an abstract of not more than 200 words.

Introduction
In a few paragraphs, state the purpose of the study and relate it to the hypothesis and the experimental design. Point out the theoretical implications of the study and relate them to previous work in the area.

1517

Next, under the subsection *Literature Review,* discuss work that had a direct impact on your study. Cite only research pertinent to a specific issue and avoid references with only tangential or general significance. Emphasize pertinent findings and relevant methodological issues. Provide the logical continuity between previous and present work. Whenever appropriate, treat controversial issues fairly. You may state that certain studies support one conclusion and others challenge or contradict it.

Method

Describe how you conducted the study. Give a brief synopsis of the method. Next develop the subsections pertaining to the *participants,* the *materials,* and the *procedure.*

Participants. Identify the number and type of participants. Specify how they were selected and how many participated in each experiment. Provide major demographic characteristics such as age, sex, geographic location, and institutional affiliation. Identify the number of experiment dropouts and the reasons they did not continue.

Materials. Describe briefly the materials used and their function in the experiment.

Procedure. Describe each step in the conduct of the research. Include the instructions to the participants, the formation of the groups, and the specific experimental manipulations.

Results

First state the results. Next describe them in sufficient detail to justify the findings. Mention all relevant results, including those that run counter to the hypothesis.

Tables and figures. Prepare tables to present exact values. Use tables sparingly. Sometimes you can present data more efficiently in a few sentences than in a table. Avoid developing tables for information already presented in other places. Prepare figures to illustrate key interactions, major interdependencies, and general comparisons. Indicate to the reader what to look for in tables and figures.

Discussion

Express your support or nonsupport for the original hypothesis. Next examine, interpret, and qualify the results and draw inferences from them. Do not repeat old statements. Create new statements that further contribute to your position and to readers understanding of it.

Conclusion

Succinctly describe the contribution of the study to the field. State how it has helped to resolve the original problem. Identify conclusions and theoretical implications that can be drawn from your study.

1518

Appendices
Place detailed information (for example, a table, lists of words, or a sample of a questionnaire) that would be distracting to read in the main body of the article in the appendices.

Notes
Use them for substantive information only, and number them serially throughout the manuscript. They all should be listed on a separate page entitled *Notes*.

References
Submit on a separate page of the manuscript a list of references with the centered heading: *References*. Arrange the entries alphabetically by surname of authors. Review the format for bibliographic entries of references in the following sample:

> Dulay, H., & Burt, M. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 16* (1), 93-95.
> Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.

List all works cited in the manuscripts in *References*, and conversely, cite all works included in *References* in the manuscript. Include in reference citations in the text of the manuscript the name of the author of the work cited, the date of the work, and when quoting, the page numbers on which the materials that you are quoting originally appeared, e.g., (Jones, 1982, pp. 235-238).

Acknowledgments
Identify colleagues who contributed to the study and assisted you in the writing process.

Author
Type the title of the article and the author's name on a separate page to ensure anonymity in the review process. Prepare an autobiographical note indicating: full name, position, department, institution, mailing address, and specialization(s). Example follows:

> JANE C. DOE, Assistant Professor, Foreign Language Education, University of America, 226 N. Madison St, Madison, WI 55306. Specializations: foreign language acquisition, curriculum studies.

Review Article

It should describe, discuss, and evaluate several publications that fall into a topical category in foreign language education. The relative significance of the publications in the context of teaching realms should be pointed out. A review article should be 15 to 20

1519

double-spaced pages.

Review

Submit reviews of textbooks, scholarly works on foreign language education, dictionaries, tests, computer software, video tapes, and other non-print materials. Point out both positive and negative aspects of the work(s) being considered. In the three to five double-spaced pages of the manuscript, give a clear but brief statement of the work's content and a critical assessment of its contribution to the profession. Keep quotations short. Do not send reviews that are merely descriptive.

Manuscripts are accepted for consideration with the understanding that they are original material and are not being considered for publication elsewhere.

Specifications for Manuscripts

All editorial correspondence, including manuscripts for publication should be sent to:

Applied Language Learning
ATFL-AP-AJ
ATTN: Editor (Dr. L. Woytak)
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

Manuscripts should be typed on one side only on 8-1/2 x 11 inch paper, double-spaced, with ample margins. Subheads should be used at reasonable intervals. Typescripts should typically run from 10 to 30 pages.

All material submitted for publication should conform to the style of the *Publication Manual of the American Psychological Association* (4th Ed., 1994) available from the American Psychological Association, P. O. Box 2710, Hyattsville, MD 20784.

Review Process

Manuscripts will be acknowledged by the editor upon receipt and subsequently sent to at least two reviewers whose area of expertise includes the subject of the manuscript. *Applied Language Learning* uses the blind review system. The names of reviewers will be published in the journal annually.

1520

Specifications for Floppy Disks

Preferably use Windows-based software. Format manuscripts produced on one of the DOS-based or Macintosh systems, as an ASQII file at double density, if possible. Please name the software used. MS Word or text documents preferred.

Copyright

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained for use in this specific issue. Requests for permission to reprint should be addressed to the Editor and should include author's permission.

1521

U.S. GOVERNMENT PRINTING OFFICE: 1999–784-149

Applied Language Learning Comment Card

We are always interested in hearing from our readers. Please use this card to comment on this issue: _____
Volume Number Date

Tear Out and Remove

MARK ONE
- ☐ New Subscription
- ☐ Address Change

PLEASE PRINT

Rank/Title First Name Middle Initial Last Name

Institution Division Department

Street City State/Country Zip Code

Tear Out and Remove

From:

_____

_____

_____

Official Business

Place
Stamp
Here

h2 51

Applied Language Learning
Defense Language Institute Foreign Language Center
Presidio of Monterey, CA 93944-5506
USA

1524

Applied Language Learning
Defense Language Institute
Foreign Language Center
Presidio of Monterey, CA 93944-5006

BULK RATE
U.S. POSTAGE
PAID
MONTEREY, CA
PERMIT NO. 42

PB-65-98-1&2
United States Army
PIN: 077 069-000
Approved for public release:
distribution is unlimited.

1525

1525